Homework 2: More Counting and Probability

For each problem, remember you must briefly explain/justify how you obtained your answer, as correct answers without an explanation will not receive full credit. Moreover, in the event of an incorrect answer, we can still try to give you partial credit based on the explanation you provide.

In general, your goal in an explanation is to write enough that a student from class who has attended lecture, but not read the problem yet, could understand your approach, verify your reasoning, and believe your answer is correct. While we do not usually need to see arithmetic, you must include enough work that in principle one could rederive your answer with only a scientific calculator.

Unless a problem states otherwise, you should leave your answer in terms of factorials, combinations, etc., for instance 26^7 or $26!/7!$ or $26 \cdot \binom{26}{7}$ are all good forms for final answers.

Instructions as to how to upload your solutions to gradescope are on the course web page.

Remember that you must tag your written problems on Gradescope.

Submission: You must upload a **pdf** of your written solutions to Gradescope under "HW 2". The use of latex is *highly recommended*. (Note that if you want to hand-write your solutions, you'll need to scan them. We will take off points for hand-written solutions that are difficult to read due to poor handwriting and neatness.)

Due Date: This assignment is due at 11:59 PM Thursday July 8 (Seattle time, i.e. GMT-7).
You will submit the written problems as a PDF to gradescope. Please put each numbered problem on its own page of the pdf (this will make selecting pages easier when you submit), and ensure that your pdfs are oriented correctly (e.g. not upside-down or sideways).

Collaboration: Please read the full collaboration policy. If you work with others (and you should!), you must still write up your solution independently and name all of your collaborators somewhere on your assignment.

1. Stuff into stuff [12 points]

(a) We have 20 (distinguishable) people and 40 (distinguishable) rooms. How many different ways are there to assign the (distinguishable) people to the (distinguishable) rooms? (Any number of people can go into any of the 40 rooms.)

(b) We have 30 identical (indistinguishable) apples. How many different ways are there to place the apples into 20 (distinguishable) boxes? (Any number of apples can go into any of the boxes.)

(c) We have 30 identical (indistinguishable) apples. How many different ways are there to place the apples into 8 (distinguishable) boxes, if each box is required to have at least two apples in it?

2. Pigeonholes and TV Shows [10 points]

Because it's the summer, you decide it's finally time to catch up on all of the shows you never got the chance to watch. Specifically, you've decided you'll spend the next 6 weeks watching all 73 episodes of Game of Thrones. You also decide to watch at least one episode every day of those 6 weeks, so you never go a day without seeing an episode. Prove that there exists a sequence of consecutive days in the next 6 weeks in which you must watch exactly 10 episodes of Game of Thrones.

When using the pigeonhole principle, be sure to mention what the pigeons and pigeonholes are.

Hint 1: Define a_i as the cumulative number of episodes you've watched up to day i for $1 \leq i \leq 42$. Now prove that $a_i = a_j + 10$ for some $i > j$.

Hint 2: $a_j + 10 \leq 83$ for all j

3. Sample Spaces and Probabilities [18 points]

For each of the following scenarios first describe the sample space and indicate how big it is (i.e., what its cardinality is) and then answer the question.

(a) You flip a fair coin 50 times. What is the probability of exactly 20 heads?

(b) You roll 2 fair 6-sided dice, one red and one blue. What is the probability that the sum of the two values showing is 10?

(c) You are given a random 5 card poker hand (selected from a single deck). What is the probability you have a full-house (3 cards of one rank and 2 cards of another rank)?

(d) 20 labeled balls are placed into 10 labeled bins (with each placement equally likely). What is the probability that bin 1 contains exactly 3 balls?

(e) There are 30 psychiatrists and 24 psychologists attending a certain conference. Three of these 54 people are randomly chosen to take part in a panel discussion. What is the probability that at least one psychologist is chosen? What is the probability that exactly three psychologists are chosen?

(f) You buy ten cupcakes choosing from 3 different types (chocolate, vanilla and caramel). Cupcakes of the same type are indistinguishable. What is the probability that you have at least one of each type?

4. Miscounting and Probabilities [12 points]

Consider the question: What is the probability of getting two three-of-a-kinds in a **7-card** poker hand (order doesn't matter). For example, this would be a valid hand: ace of hearts, ace of diamonds, ace of spaces, king of clubs, king of spades, king of hearts, and queen of clubs. (Note that a hand consisting of all 4 aces and three of the kings is also valid.)

Here is how we might compute this:

> Let E be the event where the 7 card hand has 2 three-of-a-kinds. Then, $\mathbb{P}(E) = \frac{|E|}{\binom{52}{7}}$
>
> $\binom{52}{7}$ is the total number of ways to choose a 7 card hand because there are 52 cards in the deck in total and we want to choose 7.
>
> To compute the number of hands, apply the product rule. First pick two ranks that have a 3-of-a-kind (e.g. ace and king in the example above). For the lower rank of these, pick the suits of the three cards. Then for the higher rank of these, pick the suits of the three cards. Then out of the remaining $52-6 = 46$ cards, pick one. Therefore there are
>
> $$\binom{13}{2} \cdot \binom{4}{3} \cdot \binom{4}{3} \cdot \binom{46}{1}$$
>
> hands.
>
> The probability of event E:
>
> $$\mathbb{P}(E) = \frac{\binom{13}{2} \cdot \binom{4}{3} \cdot \binom{4}{3} \cdot \binom{46}{1}}{\binom{52}{7}}$$

In this problem, you will find what is wrong with this solution.

(a) Is there overcounting of $|E|$ in the problem? That is a hand which can be produced by multiple outcomes of the sequential process. If there is, give one concrete example of such a hand and two outcomes of the process that produce it. If there is not, briefly (1-2 sentences) explain why there isn't.

(b) Is there undercounting of $|E|$ in the problem? That is a hand which cannot be produced by no outcomes of the sequential process. If there is, give one concrete example of such a hand and briefly explain why no outcome produces it. If there is no such hand, briefly explain why all hands are produced at least once.

(c) Correct the calculation – in this part you should produce a correct overall formula for the probability adjusting your $|E|$ by subtracting/dividing out any errors that would fit in (a) and adding/multiplying in any errors that would fit in (b).

(d) Optional but recommended: Find the answer differently – take a different approach to the counting in this problem (e.g. use a different sequential process). Verify that you get the same number (via a different formula) than the last part. (This question won't be graded, but it's good to think about!)

5. Guessing Game [10 points]

You take a multiple choice exam. With probability p you know the answer to the question (and get it correct). With probability $1-p$, you don't know the answer and guess randomly among the 5 possible options (of which exactly one is correct).

(a) Calculate the probability you get a question correct. Please define events and state which rules/laws you are using to do the calculation.

(b) Given that you got a question correct, what is the probability that you actually knew it (i.e., that you didn't get it correct by guessing.)? Please define events and state which rules/laws you are using to do the calculation.

6. $\mathbb{P}$NA [15 points]

Biology background: Blood Types and the Human Genome

As you may remember from basic biology, the human A/B/O blood type system is controlled by one gene for which 3 variants ("alleles") are common in the human population – unsurprisingly called A, B, and O.

As with most genes, everyone has 2 copies of this gene, one inherited from the mother and the other from the father. Everyone passes a randomly selected copy to each of their children. This happens with probability 1/2 for each copy, independently for each child. Focusing only on A and O, people with AA or AO gene pairs have type A blood; those with OO have type O blood (A is "dominant", O is "recessive").

We call the alleles that one carries their *genotype*, and the outwardly observable characteristics their *phenotype*. Thus, if a person has the genotype AA or AO, they have the phenotype A. Likewise, if they have the genotype OO, they have the phenotype O.

Notation

Please use the following notation in your answers: Let $G_I = \#\#$ be the event that person I has the genotype ##, and $Ph_I = \#$ be the event that person I has the phenotype #. For this problem, the set of possible genotypes is $\{AA, AO, OO\}$; the set of possible phenotypes is $\{A, O\}$. Use X, Y, Z, and C to refer to Xena, Yvonne, Zachary, and their Child respectively (you might not need to refer to all of them).

Give exact answers as simplified fractions and use the formulas of conditional probability to justify your reasoning for each of them (any combination of the definition of conditional probability, Bayes' Theorem, Law of Total Probability). Answers that do not explicitly use the theorems will not receive any credit. Carefully consider which theorems to use as some theorems may lead to simpler calculations than others.

The Problem

Suppose Xena and both of her parents have type A blood, but her sister Yvonne has type O.

(a) Explain what Yvonne's phenotype tells us about her and Xena's parents' genotype. With that in mind, what is the probability that Xena carries an O gene?
Hint: To start you off and to get a feel for the notation, you are calculating $\mathbb{P}(G_X = AO|Ph_X = A))$

(b) Xena marries Zachary, who has type O blood. Compute the probability that their first child will have type O blood. Make sure to represent the event using the required notation, and show all your work manipulating the equation to get your final result.

(c) If Xena and Zachary's first child had type A blood, what is the probability that Xena carries an O gene?