EECS 70 Discrete Mathematics and Probability Theory
Fall 2014 Anant Sahai **Homework 10**

This homework is due November 10, 2014, at 12:00 noon.

1. **Section Rollcall!**

In your self-grading for this question, give yourself a 10, and write down what you wrote for parts (a) and (b) below as a comment. Please put the answers in your written homework as well.

(a) What discussion did you attend on Monday last week? If you did not attend section on that day, please tell us why.

(b) What discussion did you attend on Wednesday last week? If you did not attend section on that day, please tell us why.

2. **Biased Coins Lab**

While waiting to hear back from the competitively elegant CS (Charm School) program at Cal, you decide to toss coins and plot the result to calm your nerves. Denote p as the probability of tossing a head, and $1-p$ as the probability of tossing a tail. In the previous two labs, p was 0.5, since we were exploring fair coin tosses. Now assume p can be any number such that $0 \leq p \leq 1$. As before, let k denote the total number of coin tosses.

At an abstract level, most of what this lab is about is doing the previous labs again, except that now the coin is biased. One new concept is introduced: we now need to understand how the bias of the coin affects the shapes of curves that emerge.

For each part, students who want to can choose to completely rewrite the question. Basically, you can come up with your own formulation of how to do a series of experiments that result in the same discoveries. Then, write up the results nicely using plots as appropriate to show what you observed. You can also rewrite the entire lab to take a different path through as long as they convey the key insights aimed at in each part.

Please download the IPython starter code from Piazza or the course webpage, and answer the following questions.

(a) Consider $p = 0.7, \bar{p} = 1 - p = 0.3$. If you plotted a histogram for the number of heads and tails for any number of coin tosses k, what would you expect to see that is different from the fair coin toss?

We'd expect to see that the number of heads is more than just half the number tosses. As k grows larger, this number gets closer to $0.7k$

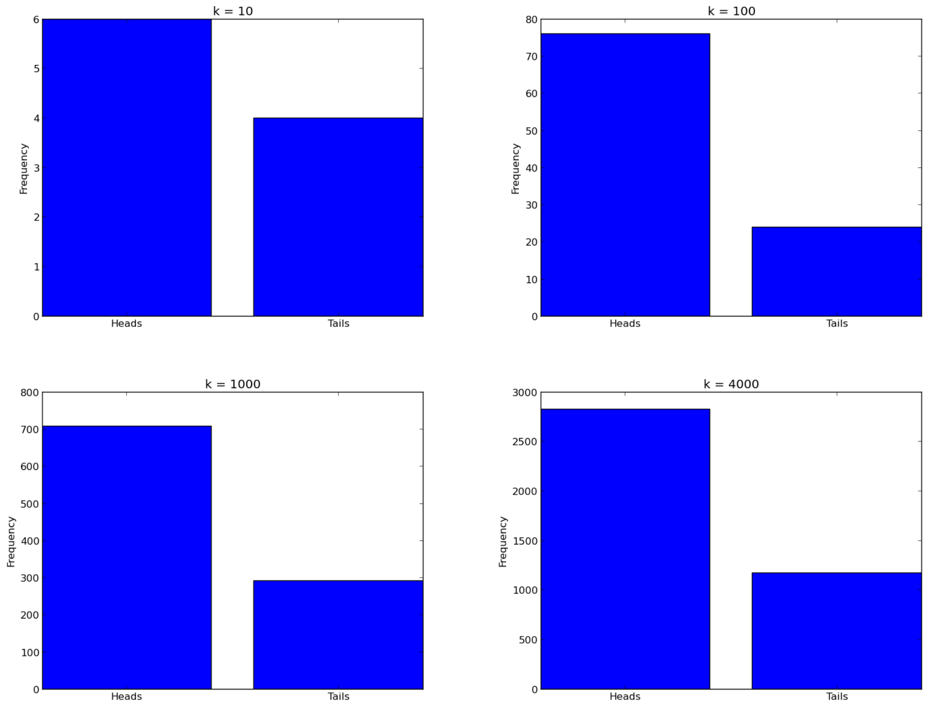
(b) If you tossed 100 coins, approximately how many heads should you get? Building on this, can you come up with an equation using $k, p, \bar{p}$ that approximately describes how many heads and tails you expect to see for any case?

Approximately 70 heads. kp, $k\overline{p}$ for heads and tails, respectively.

(c) Use a random number generator to sample a sequence of coin tosses with $p = 0.7, \bar{p} = 0.3$. Plot a bar chart for each $k = 10, 100, 1000, 4000$. Is this what you expected from parts (a) and (b)? How well does your equation match for $k = 10$? For $k = 4000$?

Hint: Implement the functions `biased_coin`, which simulates a biased coin flip with $Pr(\text{Head}) = p$, and `run_trial`, which returns the number of heads in k tosses of a biased coin with probability p of getting heads. They should be very similar to the ones you implemented previously for the fair coin.

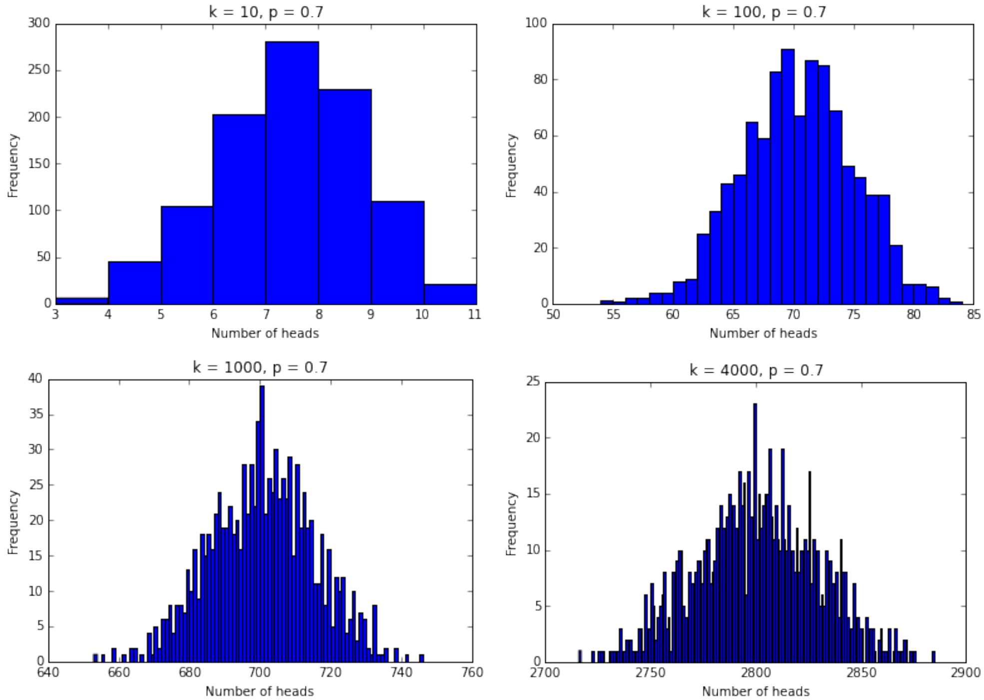
As k increases, the histograms become closer to what we predicted in previous parts.



(d) From the previous parts, you've counted the number of heads and tails for only one trial. Now fix k to be 1000. Let S_k be the total number of heads in a trial with k total coin tosses, and let m be the total number of trials.

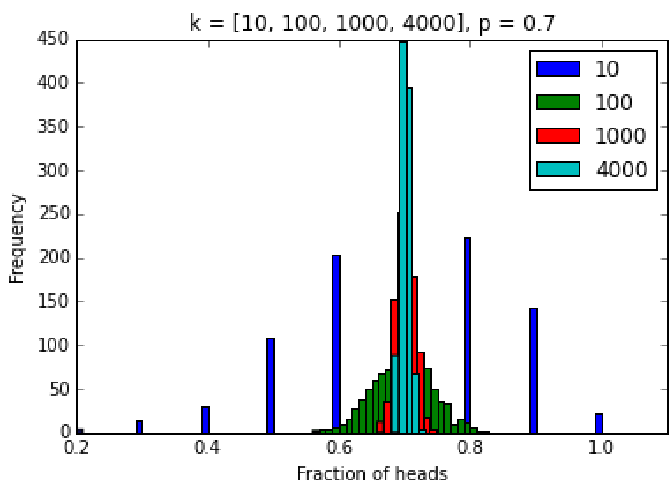
Plot a histogram of S_k for $m = 1000$ with bin size of 1. Do this again for $k = 10, 100, 4000$.

Hint: Implement the function `run_many_trials`, which returns a list of the number of heads in each trial.



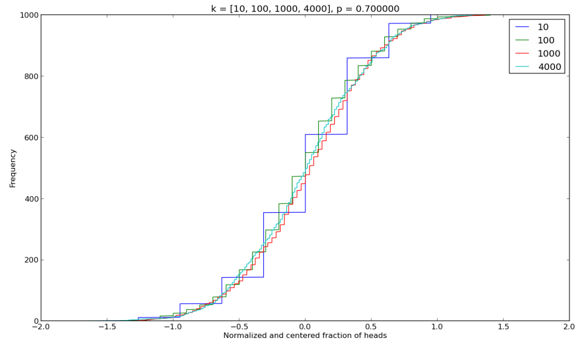
(e) Repeat the previous part, but instead plot the histograms of $\frac{S_k}{k}$ with bin size of 0.01. Make sure to plot the histograms in the same figure (so only one plot) for different values of k. How are the histograms different as k increases? Comment on what you are observing.

The mean is the same for different k values (around p), but the values spread out a lot less as we increase k.



(f) In last week's lab, we discovered a very interesting normalization by $\sqrt{k}$ that seemed to make certain curves fall right on top of each other. Let's see if that still works.

Redo the plot you did in part (g) of last week's lab, except for our biased coin and the k values you have explored above. Center the horizontal axis of the plot to be around 0.



(g) For kicks, let's try this same normalization for histograms. Plot histograms of $\frac{S_k - kp}{\sqrt{k}}$ for $m = 1000$ and $k = 10, 100, 1000, 4000$. Try 4 different bin sizes, $0.01, 0.05, 0.1, 0.5$. You should have 4 different plots, one for each bin size.

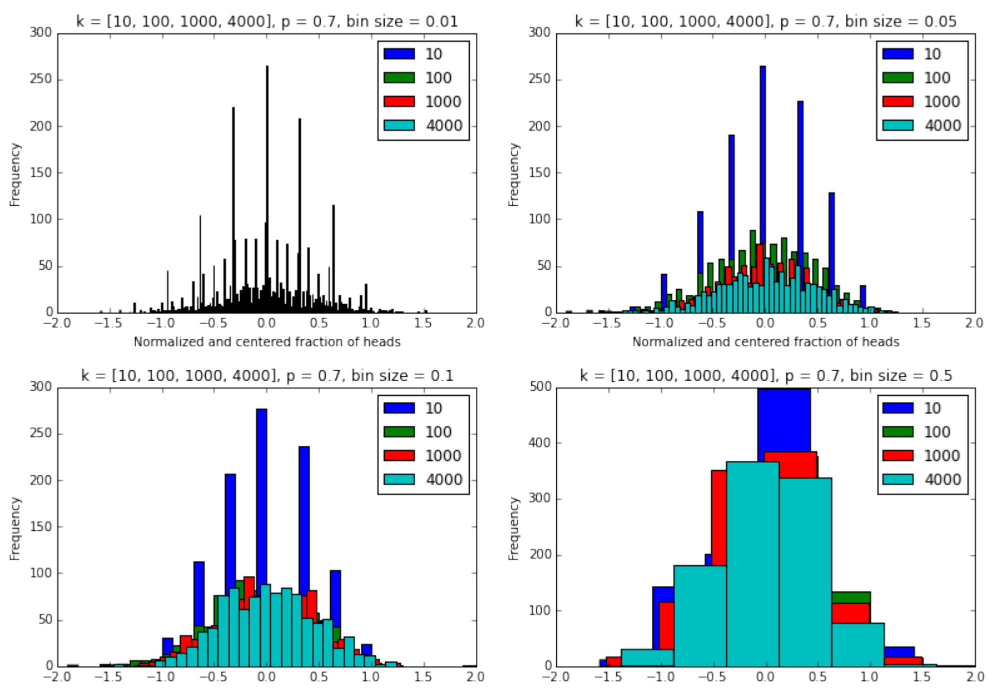
Comment on what you think the relationship is between these histograms and the "cliff-face" curves in the previous part. Think about what they mean.

The plots of the histogram align well. Their spreads are also similar now.

Looking at the equations for the plots of the "cliff-face" plot and this histogram, we're plotting the same data from different perspectives. The values we use in both are the normalized (by $\sqrt{k}$) and centered fraction of heads.

The "cliff-face" looks at the cumulative counts for the number of trials with value $\leq q$. This means we're arranging all the values in sorted order so that the n^{th} value, say, is larger than everything before it. This n^{th} value now corresponds to the normalized fraction q on the x-axis. The y-value which is n itself tells you the number of trials which resulted in a fraction less than or equal to q.

In the histogram, we're just looking at where these values are concentrated from all our different tosses.

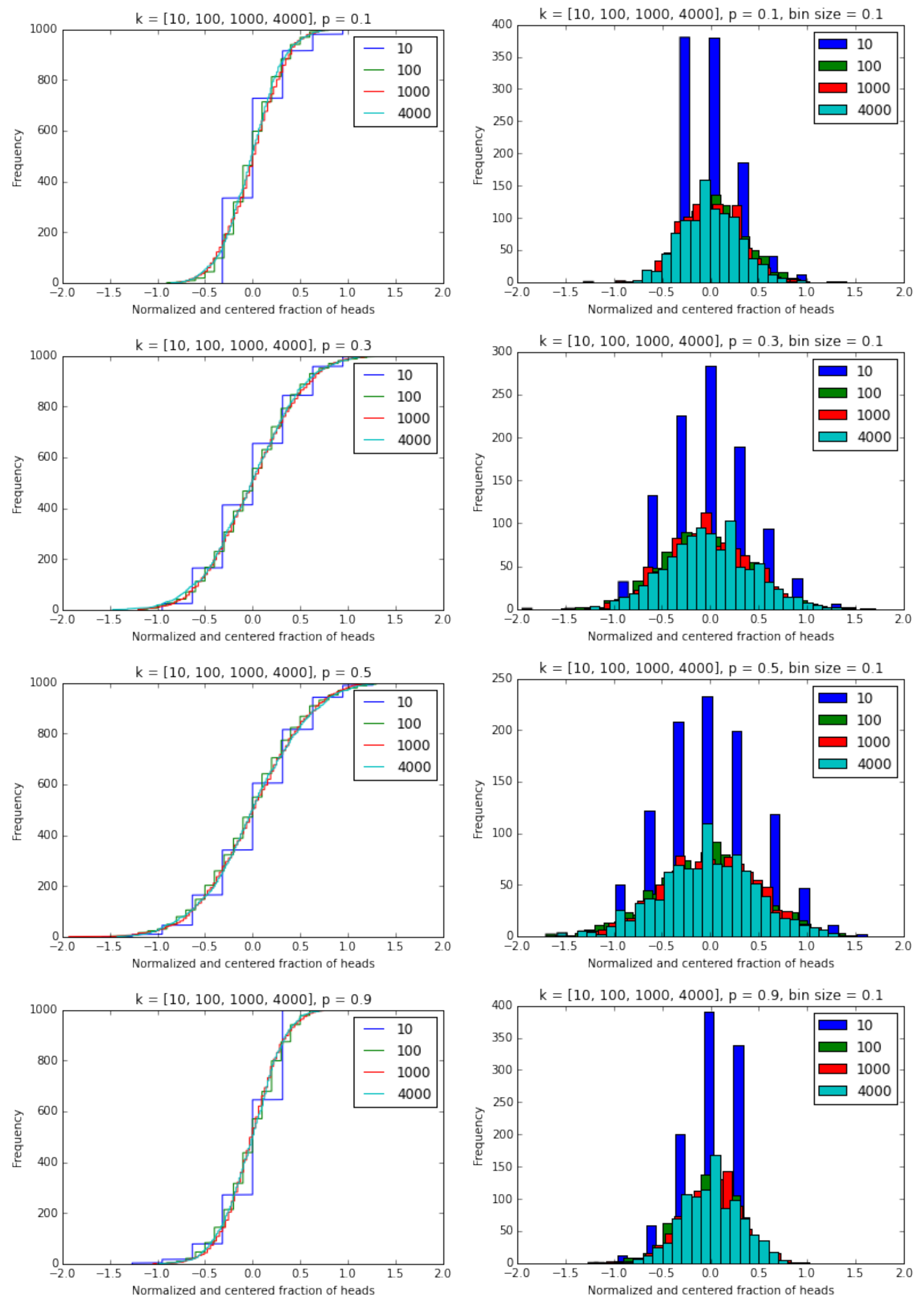


(h) Now, let's see what the effect is of varying the bias of the coin. Repeat the previous two parts for $p = 0.1, 0.3, 0.5, 0.9$. You can assume the bin size is 0.1. You should have 4 pairs of plots.

Make sure that all of these are plotted on the same scale as the previous parts. What do you observe? Which p seem more variable even after this $\sqrt{k}$ normalization? Less variable?

The spread of the plots is the largest for the case of $p = 0.5$. This makes sense intuitively because when $p = 0.5$, both heads and tails are equally likely. There is no bias towards any one of these and different sequences of heads and tails are equally probable. The results are totally random.

As the probability p deviates away from 0.5, either heads or tails are more likely and the results will reflect a bias towards one of these. And the spread of the values decreases.



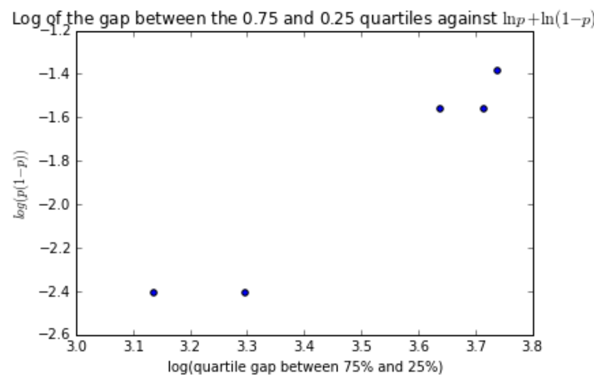
(i) Based on what you observe in the previous pattern, you decide to try and hunt out the actual dependence on the shape by looking at appropriate plots. As you did in part (f) of last week's lab, plot the Log of the gap between the 0.75 and 0.25 quartiles against $\ln p + \ln(1-p)$ in a scatter plot for $k = 4000$. (For this, you could also just plot the gap against $p(1-p)$ on a Log-Log plot.) What do you observe?

Hint: Implement the method `calculate_quartile_gap`, which calculates the gap between the first and third quartile.

You can also plot a scatter plot in Matplotlib with `plt.scatter()`.

As the log of $p(1-p)$ increases, we see that the interquartile gap also increases. Thus, for $p = 0.5$ (when $\log p(1-p)$ is the largest), the interquartile gap tends to be large too.

Again, this relates to the results being spread out more when $p = 0.5$ as both heads and tails are equally likely.

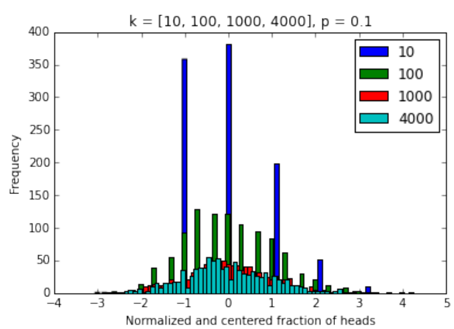


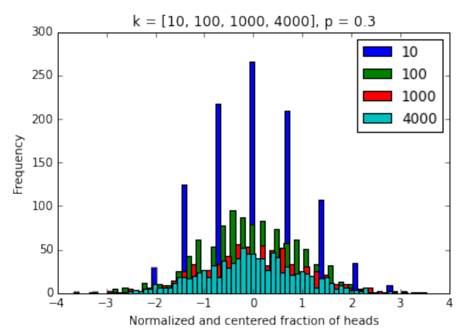
(j) This suggests another normalization. We call $\sqrt{p(1-p)}$ the "standard deviation" for a 0-1 coin toss with probability p of being 1.

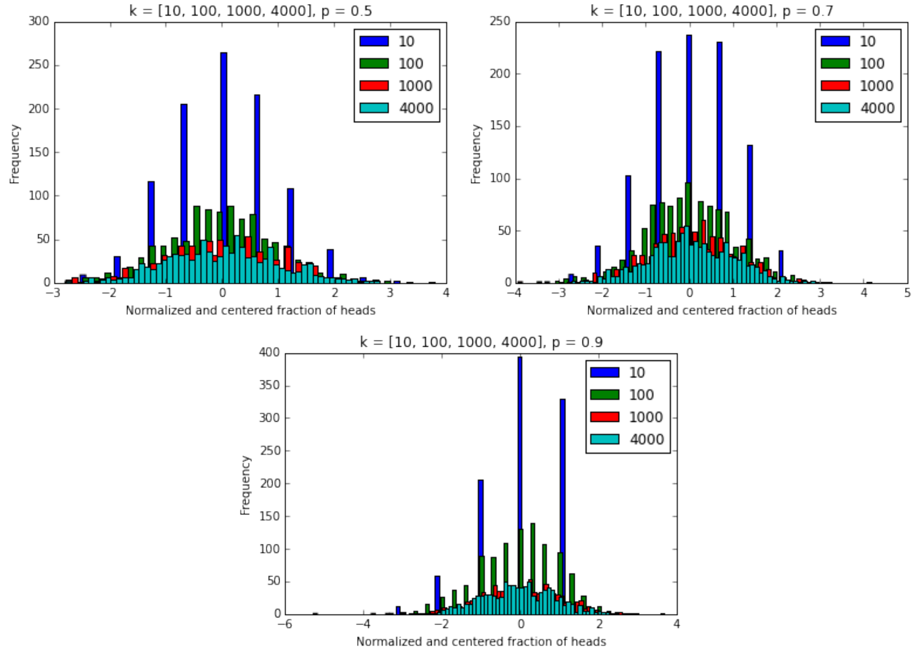
Plot a histogram of $\frac{S_k - kp}{\sqrt{k}\sqrt{p(1-p)}}$ for $k = 1000$ with bin size of 0.1. Do this again for $k = 10, 100, 4000$ and for $p = 0.1, 0.3, 0.5, 0.7, 0.9$ as before. You should have one plot for each value of p, so there should be 5 plots in total.

What do you observe?

The spread across different probabilities now is similar. It is no longer the case that $p = 0.5$ shows the most spread or that $p = 0.9$ shows a significantly smaller spread.

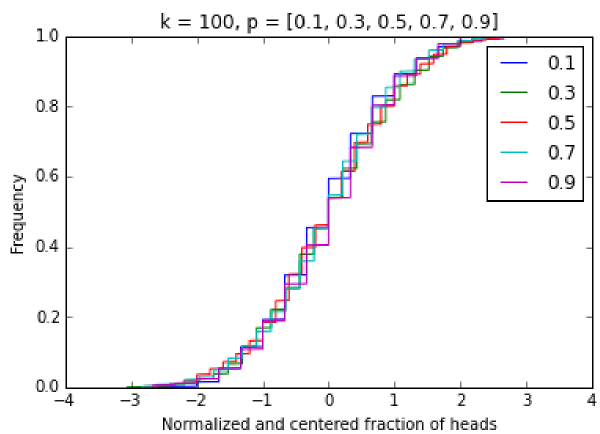


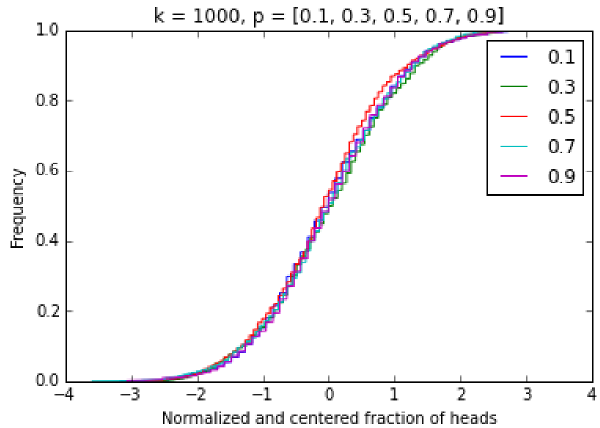


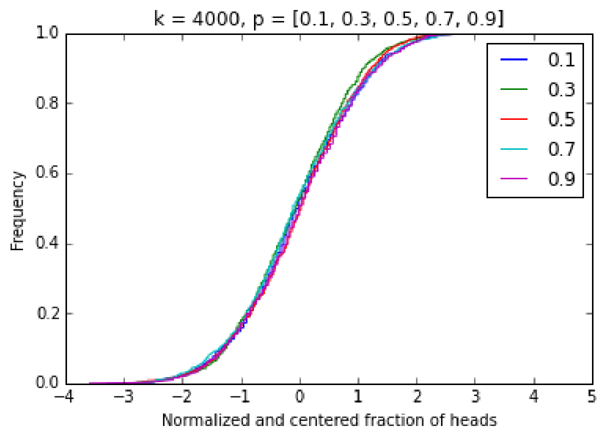


(k) Use the same normalization as the previous part and make the cliff-face plots corresponding to part (d) of last week's lab. To be precise, look at the range $d = -3$ to $d = +3$ and plot how often (out of the m runs — as a fraction between 0 and 1) $\frac{S_k - kp}{\sqrt{k}\sqrt{p(1-p)}}$ is less than d.

This should be an increasing curve. Here, put all the different p plots together but have three different plots for $k = 100, 1000, 4000$. (You already know from earlier plots that the different k's track each other.)





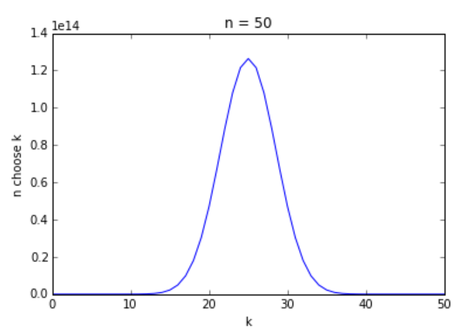


(l) Lastly, we would like to explore a function that defines how many ways you can choose k distinct objects out of n possible objects. This is written as $\binom{n}{k}$ and is read aloud as "n choose k". We define $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. For example, $\binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(3 \cdot 2 \cdot 1) \cdot (2 \cdot 1)} = \frac{120}{12} = 10$. Plot the value $\binom{50}{k}$ on the y-axis and k on the x-axis for $0 \leq k \leq 50$.

Does this value constantly grow as k gets larger? What does the shape of the graph remind you of?

Hint: Implement the function `choose(n, k)` using `math.factorial`.

The value of $\binom{n}{k}$ does *not* always get larger as k gets larger, it actually peaks around $k = \frac{n}{2}$ and is symmetric around this value. Looking at the plot below, the shape looks very similar to a bell-shaped curve.



Reminder: When you finish, don't forget to convert the notebook to pdf and merge it with your written homework. Please also zip the `ipynb` file and submit it as `hw10.zip`.

3. Charm School Applications

(a) n males and n females apply to the Elegant Etiquette Charm School (EECS) within UC Berkeley. The EECS department only has n seats available. In how many ways can it admit students? Use the above story for a combinatorial argument to prove the following identity:

$$\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2$$

One way of counting is simply $\binom{2n}{n}$, since we must pick n students from $2n$.
The other way is to first pick i males, then $n-i$ females. Equivalently, choose i males to admit, and i females to NOT admit. For a fixed i, this yields $\binom{n}{i}\binom{n}{n-i} = \binom{n}{i}^2$ choices. Thus, over all choices of i:

$$\binom{2n}{n} = \sum_{i=0}^{n} \binom{n}{i}^2$$

(b) Among the n admitted students, there is at least one male and at least one female. On the first day, the admitted students decide to carpool to school. The boy(s) get in one car, and the girl(s) get in another car. Use the above story for a combinatorial argument to prove the following identity:

$$\sum_{k=1}^{n-1} k \cdot (n-k) \cdot \binom{n}{k}^2 = n^2 \cdot \binom{2n-2}{n-2}$$

(Hint: Each car has a driver...)

Out of the n males and n females who applied, count the number of ways that accepted students can drive to school.

RHS: First pick one male driver and one female driver from the n male and n female applicants (n^2). Then pick the other $n-2$ accepted students from the pool of $2n-2$ remaining applicants.

LHS: Pick k males and $n-k$ females that were accepted: $\binom{n}{k}\binom{n}{n-k} = \binom{n}{k}^2$. Then pick a driver among the k males, and among the $n-k$ females. Because the problem statement says there is at least 1 girl and 1 boy, k can range from 1 to $n-1$.

4. Getting to CS

Harry, the chosen one, is chosen to drive the boys to the Charm School. The Flying Ford Anglia he's driving is behaving weirdly – it would only go south or east for at least a certain distance. Figure 1 shows the path the car could go from Harry's house (H) to the Charm School (CS).

(a) How many ways can he get there?

Harry needs to go south 6 times and east 7 times. We can think of a path from H to CS as a series of $6+7=13$ moves, 6 of which are going south and 7 of which are going east. Out of 13 total moves, there are $\binom{13}{6} = \binom{13}{7} = 1,716$ ways to make such a path. (Either choose 6 out of 13 moves to be going south, or choose 7 out of 13 moves to be going east.) $\square$

(b) Harry has to pick up other students at point P. How many ways can he stop by point P and go to the Charm School?

Let $P_{x_1 \to x_2 \to \ldots \to x_n}$ denote the set of paths from x_1 to x_n that pass through $x_2, x_3, \ldots, x_{n-1}$. Then the number of ways Harry can get to CS by passing P is just the number of ways from H to P multiplied by the number of ways from P to CS.

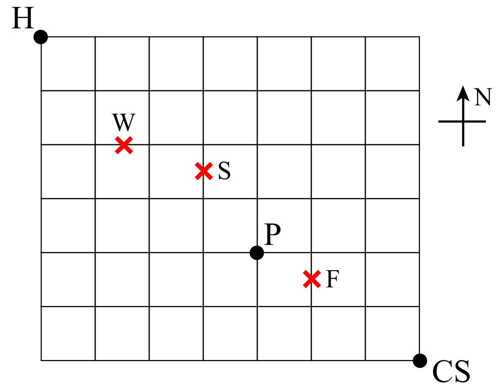


Figure 1: The map from Harry's house to the Charm School

$$|P_{H\to P\to CS}| = |P_{H\to P}||P_{P\to CS}| = \binom{4+4}{4}\binom{2+3}{3} = \binom{8}{4}\binom{5}{3} = 70\cdot 10 = 700.$$

□

(c) The Whomping Willow (W) will attack anything that comes near. Harry must not drive through it. How many ways can he pick up the students and then go to the Charm School now?

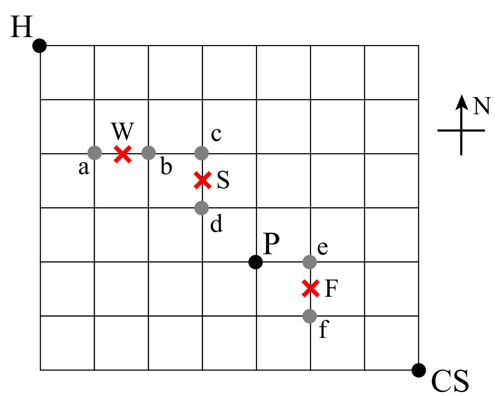


Figure 2: More elaborated map from Harry's house to the Charm School

Figure 2 labels more points. Notice that point W doesn't affect the paths from P to CS at all, so we can just calculate how it affects the number of possible paths from H to P, then multiply with $|P_{P\to CS}|$. First, we find the total number of paths that pass through W. For any path to pass through W, it must pass point a then point b.

$$|P_{H\to a\to b\to P}| = |P_{H\to a}||P_{b\to P}| = \binom{3}{1}\binom{4}{2} = 3\cdot 6 = 18$$

The number of paths from H to P that don't pass W is just the total number of paths from H to P minus the number of paths that do pass W.

$$|P_{H\to P}| - |P_{H\to a\to b\to P}| = 70 - 18 = 52.$$

Therefore, the total number of paths from H to P to CS that don't pass $W = 52 \cdot |P_{P\to CS}| = 52 \cdot \binom{5}{3} = 52 \cdot 10 = 520$. □

(d) On top of the Whomping Willow, the Marauder's Map shows Professor Snape (S) and Filch (F) whom he doesn't want to drive past either. How many ways can Harry go to the Charm School without getting past the Whomping Willow (W), Professor Snape (S), or Filch (F), while still picking up other students?

Again, the path that passes S must pass through points c and d, and the path that passes F must pass through points e and f.

$$|P_{H\to c\to d\to P}| = \binom{5}{3}\binom{2}{1} = 20$$
$$|P_{P\to e\to f\to CS}| = \binom{3}{1} = 3$$

Observe that points W and S only affect the paths from H to P, and F only affects the paths from P to CS. We can calculate these separately.

<u>From H to P:</u> It is a little tricky to find the number of paths that don't pass through both W and S. First, we find the total number of paths that pass through both W and S,

$$|P_{H\to a\to b\to c\to d\to P}| = \binom{3}{1}\binom{2}{1} = 6$$

Next, we find the total number of paths that passes *either W or S*,

$$\begin{aligned}|P_{H\to a\to b\to P} \cup P_{H\to c\to d\to P}| &= |P_{H\to a\to b\to P}| + |P_{H\to c\to d\to P}| - |P_{H\to a\to b\to c\to d\to P}| \\ &= 18 + 20 - 6 = 32.\end{aligned}$$

Hence, the number of paths that don't pass W and S, i.e., any of the points a, b, c, and d, is,

$$\begin{aligned}|P_{H\to P}| - |(P_{H\to a\to b\to P} \cup P_{H\to c\to d\to P})| &= |P_{H\to P}| - |P_{H\to a\to b\to P} \cup P_{H\to c\to d\to P}| \\ &= 70 - 32 = 38\end{aligned}$$

<u>From P to CS:</u> The number of paths that don't pass F is $|P_{P\to CS}| - |P_{P\to e\to f\to CS}| = \binom{5}{3} - 3 = 7$.

Therefore, the answer is

$$(|P_{H\to P}| - |(P_{H\to a\to b\to P} \cup P_{H\to c\to d\to P})|)\,(|P_{P\to CS}| - |P_{P\to e\to f\to CS}|) = 38 \cdot 7 = 266.$$

□

5. Handshakes

In the first day of Charm School, all students are going to practice self-introduction skills. When they meet a new person, they will introduce themselves and shake their hands. Suppose that there are m students who are right-handers and n students who are left-handers. When two people are both right-handers, they will shake their right hands and so will left-handers. When a right-hander meets a left-hander, they will either both shake their right hands or both shake their left hands, with equal probability $p = 0.5$.

(a) Assume that all students do not know each other at the beginning. How many handshakes will occur?

There are a total of $m + n$ students and any 2 of them will shake their hands with each other. Therefore the number of handshakes will be $\binom{m+n}{2}$.

(b) Assume that there are k people who have known each other so they will not do any handshake. How many handshakes will occur?

This means that k of the n people will not shake their hands with each other, so $\binom{k}{2}$ handshakes will not occur among them. The total number of handshakes is therefore $\binom{m+n}{2} - \binom{k}{2}$.

(c) From now on, we assume all students do not know each other at the beginning. We know that one of the students, Ginny, is a right-hander. What is the probability that Ginny will shake her right hand when she meets the first person?

Let A be the event that Ginny will shake her right hand when she meets the first person, R be the event that the first person she meets is a right-hander, and L be the event that the first person she meets is a left-hander. Then we have by the law of total probability

$$P(A) = P(A|R)P(R) + P(A|L)(L).$$

We know that $P(R) = \frac{m-1}{m+n-1}$ and $P(L) = \frac{n}{m+n-1}$. Since Ginny is a right-hander, when she meets a right-hander, they will always shake their right hands, and hence $P(A|R) = 1$. When she meets a left-hander, there will be 0.5 probability that they will shake their right hands, and hence $P(A|L) = \frac{1}{2}$. So

$$P(A) = 1 \times \frac{m-1}{m+n-1} + \frac{1}{2} \times \frac{n}{m+n-1} = \frac{m + \frac{1}{2}n - 1}{m+n-1}.$$

(d) If we know that Neville shakes his left-hand in his first handshake, what is the probability that Neville is a left-hander?

Let A be the event that Neville shakes his left-hand in his first handshake, L be the event that Neville is a left-hander and R be the event that Neville is a right-hander. By Bayes' Rule we have:

$$P(L|A) = \frac{P(A|L)P(L)}{P(A|L)P(L) + P(A|R)P(R)}$$

We know that $P(L) = \frac{n}{m+n}$ and $P(R) = \frac{m}{m+n}$. Similar to (c), we know that given Neville is a left-hander, the probability that he will shake his right hand in his first handshake is: $P(A|L) = \frac{n-1}{m+n-1} + \frac{1}{2}\frac{m}{m+n-1}$. If Neville is a right-hander, the only possible situation that he shakes his left hand in his first handshake is if he meets a left-hander (with probability $\frac{n}{m+n-1}$), and they shake their left hands (with probability $\frac{1}{2}$). Therefore $P(A|R) = \frac{1}{2}\frac{n}{m+n-1}$.

Then we can get

$$P(L|A) = \frac{\left(\frac{n-1}{m+n-1} + \frac{1}{2}\frac{m}{m+n-1}\right) \times \frac{n}{m+n}}{\left(\frac{n-1}{m+n-1} + \frac{1}{2}\frac{m}{m+n-1}\right) \times \frac{n}{m+n} + \frac{1}{2}\frac{n}{m+n-1} \times \frac{m}{m+n}} = \frac{\frac{1}{2}m + n - 1}{m+n-1}.$$

(e) Let x be the number of right-handshakes occurring during the Charm School. What is the possible range of x?

There are m right-handers, so there will be at least $\binom{m}{2}$ right-handshakes, that is $x \geq \binom{m}{2}$. More right-handshakes will occur when right-handers meet left-handers. There are a total of at most mn handshakes among right-handers and left-handers, so the number of right-handshakes should not be more than $\binom{m}{2} + mn$. The possible range is:

$$\binom{m}{2} \leq x \leq \binom{m}{2} + mn.$$

It is also correct if your upper bound is $\binom{m+n}{2} - \binom{n}{2}$, which is the total number of handshakes minus the minimum number of left-handshakes. They are just different approaches to count the same thing and therefore they are equal.

$$\begin{aligned}\binom{m+n}{2} - \binom{n}{2} &= \frac{(m+n)(m+n-1)}{2} - \frac{(n)(n-1)}{2} \\ &= \frac{m(m+n-1)+n(m+n-1)-n(n-1)}{2} \\ &= \frac{m(m-1)+mn+nm+n(n-1)-n(n-1)}{2} \\ &= \frac{m(m-1)+2mn}{2} \\ &= \frac{m(m-1)}{2}+mn \\ &= \binom{m}{2}+mn\end{aligned}$$

(f) If x is within the range you provided in (e), what is the probability that there are exactly x right-handshakes?

Since $\binom{m}{2}$ handshakes must occur among the m right-handers, so we just need to consider the remaining $x - \binom{m}{2}$ right-handshakes that occur among the m right-handers and n left-handers. Let $k = x - \binom{m}{2}$. Among the mn handshakes between right-handers and left-handers, there are $\binom{mn}{k}$ possible outcomes that result in k right-handshakes. Since the probability of a left-handshake occurring and a right-handshake occurring are both $\frac{1}{2}$, the probability of each possible outcome is $(\frac{1}{2})^k \times (\frac{1}{2})^{mn-k} = (\frac{1}{2})^{mn}$. This is analogous to tossing k heads out of n fair coins. Therefore, the probability that there are exactly x right-handshakes is :

$$(\frac{1}{2})^{mn}\binom{mn}{k} = (\frac{1}{2})^{mn}\binom{mn}{x-\binom{m}{2}}.$$

(g) If we randomly pick $y \leq \min\{m,n\}$ students, what is the probability that there are exactly z students who are right-handers? (Assume $0 \leq z \leq y$.)

The number of combinations of choosing y students is $\binom{m+n}{y}$. The number of combinations of choosing y students that include exactly z right-handers is the number of ways of choosing z right-handers among the m students multiplied by the number of ways of choosing the remaining $y-z$ left-handers among the n students. This is just $\binom{m}{z}\binom{n}{y-z}$. So the probability that there are exactly z students who are right-handers is:

$$\frac{\binom{m}{z}\binom{n}{y-z}}{\binom{m+n}{y}}.$$

(h) Let $m = 5$ and $n = 5$. If we randomly pick $y = 4$ students among them, what is the probability that there are exactly 5 right-handshakes among the 4 students?

Let A be the event that there are exactly 5 right-handshakes among the 4 students, and B_i be the event that there are i persons who are right-handers among the 4 students. By the total law of probability we have:

$$P[A] = P[A|B_0]P[B_0] + P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + P[A|B_3]P[B_3] + P[A|B_4]P[B_4]$$

From (g), we know that for $0 \le i \le 4$, we have:

$$P[B_i] = \frac{\binom{m}{i}\binom{n}{y-i}}{\binom{m+n}{y}} = \frac{\binom{5}{i}\binom{5}{4-i}}{\binom{10}{4}}.$$

So, $P[B_0] = \frac{5}{210}$, $P[B_1] = \frac{50}{210}$, $P[B_2] = \frac{100}{210}$, $P[B_3] = \frac{50}{210}$ and $P[B_4] = \frac{5}{210}$.
If there is no right-hander or just one right-hander, it is not possible to have 5 right-handshakes. Therefore $P[A|B_0] = P[A|B_1] = 0$. If all of the 4 students are right-handers, there will be at least $\binom{4}{2} = 6$ right-handshakes. Therefore $P[A|B_4] = 0$. Similar to (f), we know that for $2 \le i \le 3$, we have:

$$P[A|B_i] = (\frac{1}{2})^{i(4-i)} \binom{i(4-i)}{5-\binom{i}{2}}.$$

So $P[A|B_2] = \frac{1}{16}$ and $P[A|B_3] = \frac{3}{8}$.
Then

$$P[A] = \frac{1}{16} \times \frac{100}{210} + \frac{3}{8} \times \frac{50}{210} = \frac{5}{42}.$$

6. **Seating Arrangements**
The Charm School features full course dinner. Everyone is assigned a table. Each table is a round table with ten seats. Each seat has a nametag.

(a) The staff forgot to put up the name tags on table 1. How many ways can students at that table sit? (Note that ABCDEFGHIJ and BCDEFGHIJA are the same seating arrangement because the table is round, and all seats are identical.)

If we look at the table from the top, read the order of students from one particular seat clockwise, we will get a string of ten student names such as ABCDEFGHIJ. Notice that these 10 patterns, ABCDEFGHIJ, JABCDEFGHI, IJABCDEFGH, ..., and BCDEFGHIJA, are the same because A just moves one seat clockwise.

Since there are 10! ways to arrange ten students' names in a straight line, there are $\frac{10!}{10} = 9!$ ways to sit them around a round table. □

(b) Table 2 has name tags but all 10 students there didn't see the name tags and just sat randomly. If all students have different names, what is the probability that all of them have correct name tags?

Each seat is different now. So the total number of ways to sit is 10! (10 students can sit in first student's seat, 9 students can sit in the second student's seat, etc.). There is only one way for them to have all correct name tags so the probability is $\frac{1}{10!}$. □

(c) What are the chances all the students in table 2 have wrong name tags? (*Hint: Try small cases first and see if you can form a recurrence relation. You might want to use a calculator or ask your friend Wolfram to calculate the final result.*)

Let D_n denote the number of ways to sit n people such that none of them is in their seat. Let s_i denote the i^{th} student arrived. Consider small cases first.

- <u>$n = 2$:</u> s_1 can only sit on s_2's seat, and vice versa. Therefore, $D_2 = 1$.
- <u>$n = 3$:</u> If s_2 sits on s_1's seat, then s_1 must be on s_3's seat, otherwise s_3 will have to sit in their own seat. Similarly, if s_3 sits on s_1's seat, there is only one way to sit s_1 and s_2. Therefore, $D_3 = 2$.

- $\underline{n = 4}$: If s_2 sits on s_1's seat, then s_1 either sits on s_2's seat or doesn't. If s_1 sits on s_2's seat then s_3 and s_4 are left with their own seats. We know from $n = 2$ that there is $D_2 = 1$ way to seat them. If s_1 doesn't sit on s_2's seat, then there are 3 students, each of which has a unique seat they can't sit on. We also know from $n = 3$ that there are $D_3 = 2$ ways to seat them. Therefore, there are $D_2 + D_3 = 1 + 2$ ways for all people to not sit at their seats when s_2 is on s_1's seat. Since there are 3 students who can sit on s_1 seats (s_3 and s_4 in addition to s_2), the total number of ways is $3(D_2 + D_3) = 3 \cdot 3 = 9$.

Now we can generalize to D_n. Let s_x be the student who sits on s_1's seat. If s_1 sits on s_x's seat, we have $n-2$ people left with their respective seats, and we know they can sit in D_{n-2} ways. Otherwise, treat s_x's seat as s_1's seat and we have $n-1$ people with their respective seats, and there are D_{n-1} ways to sit them. Since $n-1$ students can sit on s_1's seat, the total number of ways n students can all sit in the wrong seat is

$$D_n = (n-1)(D_{n-1} + D_{n-2}). \tag{1}$$

Next, we find D_{10} from Equation (1), starting with $D_2 = 1$ and $D_3 = 2$, as shown in Table 1.

n	D_n
2	1
3	2
4	9
5	44
6	265
7	1,854
8	14,833
9	133,496
10	1,334,961

Table 1: Total number of ways n students can all sit in wrong seats

Hence, the probability that all students have the wrong name tags is $\frac{D_{10}}{10!} = \frac{1,334,961}{10!} = 0.3679$. □

Note: This D_n has a name. We call a permutation such that no item is in its original position a *derangement*. In general it is denoted as $!n$.

(d) Table 3 has two Freds and three Georges. The rest have distinct names. If all the students there sit randomly, what is the probability that they all have correct name tags?

Total number of ways to sit is still 10!, but the number of correct ways increases to 2!3! (because the two Freds can permute and still have correct name tags, so can the three Georges). So the probability that they all have correct name tags is $\frac{2!3!}{10!}$. □

(e) At table 10, only the first student arrived didn't see the name tag, so he sat randomly. The 2nd student to arrive sits on her seat if it is free, otherwise she sits on a random seat. The 3rd student to arrive sits on his seat if it is free, otherwise he sits on a random seat. And so on. If all students have different names, what is the probability the last student gets to sit at his/her seat? (*Hint: Again, try small cases first and observe what's going on.*)

Again, let s_i denote the i^{th} student arrived. Let n denote the number of students and seats. Trying small cases should give you the intuition that *the chances the last student, s_n, gets to sit in his/her own seat is 1/2, regardless of n*. Why? Because the only decision of s_1 that matters is between his/her own seat and s_n's seat. If s_1 decides to take other student's seat, say s_x's, then the same situation happens again when it's s_x's turn to sit. (Everybody else will sit in their own seat unless it's taken, and s_x is just the new s_1.) Now s_n's fate depends on s_x if and only if s_x chooses his/her own seat or s_n's seat, otherwise the decision falls upon another student, s_y, whose seat has been taken by s_x, and so on. Since there are 2 seats to choose between, the probability s_n gets to sit in his/her own seat is 1/2. Let us prove this properly.

Claim: For n students with n seats, the probability the last student gets to sit on his/her seat is 1/2.

Proof: We will prove by strong induction.
Base Case: $n = 2$. The probability s_2 sits in his/her own seat = the probability s_1 sits in his own seat = 1/2. Thus, the base case is true.
Inductive Hypothesis: Assume the statement is true for all $2 \le n \le k$.
Inductive Step: We have to show the statement is true for $n = k+1$. Let L denote the event s_n sits in his/her own seat. There are 3 cases.

- s_1 takes his/her own seat. Call this event A. $P(A) = 1/(k+1)$. Everybody can sit in their own seat, including s_n, so $P(L|A) = 1$.
- s_1 takes s_n's seat. Call this event B. $P(B) = 1/(k+1)$. The last student cannot sit in his/her own seat no matter how other students sit, so $P(L|B) = 0$.
- s_1 takes other student's seat. Call this event C. $P(C) = (k-1)/(k+1)$. Let s_x be the student whose seat is taken by s_1. $s_2, s_3, \ldots, s_{x-1}$ will sit in their own seat. When s_x arrives, there will be $y = n - x + 1 \le k$ seats left, with y students to seat. s_x gets to sit first and s/he picks the seat randomly. This is similar to the case with y students since the beginning. Because $y \le k$, we know by the Inductive Hypothesis that s_n will get to sit on his/her seat with probability 1/2. So, $P(L|C) = 1/2$.

Since events A, B, and C partition the sample space, we can find $P(L)$ using the Law of Total Probability,

$$\begin{aligned} P(L) &= P(L|A)P(A) + P(L|B)P(B) + P(L|C)P(C) \\ &= 1\left(\frac{1}{k+1}\right) + 0\left(\frac{1}{k+1}\right) + \frac{1}{2} \cdot \frac{k-1}{k+1} \\ &= \frac{2+k-1}{2(k+1)} \\ &= \frac{1}{2}. \end{aligned}$$

This completes the proof. □

7. Dinner Time

Now let's move on to the actual dinner. Each person has all sorts of plates, flatwares, and glasses in front of them, as shown in Figure 3.[1] The basic rule is to start using utensils furthest from your plate and end with the closest ones. Table 2 lists the menu and the corresponding utensils.

(a) Ron is confused what utensils to use (*'Wait, I think I'm at the wrong Charm School..'*). Fortunately, he can wait for his server to select the right plates and glasses. He just needs to pick flatwares. All he sees are, 4 forks (B, C, D, and N), 3 knives (H, I, and L), and 2 spoons (J and M). So, for each course served, he mimics what other people are using. For example, if other people are using a fork

[1] Source: `http://damoneroberts.tumblr.com/post/51078389219/home-tip-of-the-day-proper-place-settin`

Menu	Plates	Flatwares	Glasses
Water	-	-	O
Red wine	-	-	P
White wine	-	-	Q
Bread	K	L	-
Soup	-	J	-
Salad	E	B	-
Fish	F	C, I	-
Meat	G	D, H	-
Dessert	-	M, N	-

Table 2: Courses and utensils

Figure 3: Formal dinner setting

and a knife, he picks one fork and one knife. (He can't tell the difference between each fork, but can separate forks from knives and spoons just fine.) What is the probability he uses all utensils correctly? Each utensil is collected after each course and can't be used twice.

We can think of this as Ron mentally making 3 separate orderings for forks, knives, and spoons. For each plate he finds out what kind(s) of flatware to use, then pick them according to the orders in his 3 lists. For example, let his lists be, forks: D, C, N, B, knives: I, H, L, and spoons: M, J, then he picks I for bread, M for soup, D for salad, C and H for fish, N and L for meat, and B and J for dessert. See Table 3 for illustration.

There are 4! ways to order the forks, 3! ways to order the knives, and 2! ways to order the spoons. So there are 4!3!2! ways to make these 3 lists. Since there is only one correct way to pick all the flatwares, the probability he uses all utensils correctly is $\frac{1}{4!3!2!} = \frac{1}{288}$. □

(b) Luna just doesn't care. For each course she just picks one or two random flatwares so that all of them are used at the end, and forces the server to serve on one random plate. For each drink she picks a random glass. What is the probability she used at least two things wrong? (If a utensil isn't used in the course it is matched with, then it is used wrongly.)

Again, we can view choosing utensils for each meal as just virtually arranging and using them in order.

Menu	Forks	Knives	Spoons
Bread	-	I	-
Soup	-	-	M
Salad	D	-	-
Fish	C	H	-
Meat	N	L	-
Dessert	B	-	J

Table 3: How Ron can use flatwares according to the example orderings (in columns).

There are 4! and 3! ways to arrange plates and glasses. For flatwares, it is trickier.

Luna can use 1 or 2 per dish. Since there are 9 of them and 6 dishes to use with, they must be partitioned into 3 groups of one and 3 groups of two. First, we permute all flatwares and make Luna use the first three individually, and the last 6 in pairs (or a partition of 111222). This gives $\frac{9!}{2!2!2!}$ ways to arrange them since we permute 9 flatware items but the order does not matter for those grouped in pairs. Now, the partition 111222 can also be permuted (111222, 121122, 121212, … so on), with a total number of $\frac{6!}{3!3!}$ or $\binom{6}{3}$ ways. Therefore, the total number of ways Luna can choose for all flatwares is

$$\frac{9!6!}{2!2!2!3!3!} \times 4! \times 3! = \frac{9!6!}{2!} = 1.306 \times 10^8.$$

The probability of using at least two things wrong = 1 − probability of using at most one thing wrong. There is only one way to use everything correctly. The only place Luna can use just one thing wrong is with the flatwares where one of the paired flatwares is used with a flatware that is supposed to be used alone, otherwise at least two things must be wrong (because the other utensil whose place has been taken must also be wrong). For a pair of flatwares, there are 2 ways to pick just one of them, and 3 flatwares that it can be paired with (those that are supposed to be used alone). There are 3 pairs of flatwares, so the total number of ways to use one thing wrong = $2 \times 3 \times 3 = 18$.

Therefore, the probability that Luna used at least two things wrong = $1 - \dfrac{1+18}{9!6!/2} \approx 1$. □

(c) (*Optional*) What is the probability Hermione used all correct plates, flatwares, and glasses?

1, duh. □

8. Charming Star

At the end of each day, students will vote for the most charming student. There are 5 candidates and 100 voters. Each voter can only vote once, and all of their votes weigh the same.

(a) How many possible voting combinations are there for the 5 candidates?

Let x_i be the number of votes of the i-th candidates. We would like to find all possible combinations of $(x_1, x_2, x_3, x_4, x_5)$ such that

$$x_1 + x_2 + x_3 + x_4 + x_5 = 100.$$

It is equivalent to selecting $k = 100$ objects from $n = 5$ categories. The number of possible combinations is:

$$\binom{100+5-1}{100} = \binom{104}{100} = 4598126.$$

(b) How many possible voting combinations are there such that exactly one candidate gets more than 50 votes?

Now we have a constraint that one of the x_i should be at least 51. Say, let x_1 be at least 51. It is equivalent to giving the first candidate 51 votes at the beginning and then distributing the remaining 49 votes to them again. The number of possible combinations is $\binom{49+5-1}{49}$. Since one of the 5 candidates could have at least 51 votes, the total number of possible voting combinations such that exactly one candidate gets more than 50 votes is:

$$\binom{5}{1}\binom{49+5-1}{49} = 1464125.$$

9. Write your own problem

Write your own problem related to this week's material and solve it. You may still work in groups to brainstorm problems, but each student should submit a unique problem. What is the problem? How to formulate it? How to solve it? What is the solution?

10. Bonus

As you have noticed, this homework is themed around the idea of charm school. Bonus points for anyone who comes up with a problem related both to a real-world lesson about etiquette and charm, as well as to counting and/or basic probability. Is there something that you wish your fellow students knew better as far as charm and etiquette goes? Everything is fair game: hygiene, dress, grooming, manners, conversation, politeness, caring, formality, hospitality, open-mindedness, networking, etc. Help make EECS a more charming, welcoming, and gracious environment...