EECS 70 Discrete Mathematics and Probability Theory
Fall 2014 Anant Sahai Homework 11

This homework is due November 17, 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. 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, Birthday Paradox, and Stirling's Approximation Lab**

   Up until this point, everything that you have done in the last three virtual labs is something that you could've naturally discovered yourself as something worth trying. The data is speaking directly to the experimentalist in you. However, discovering an actual formula for the shape of this "cliff-face" is something that actually requires a theoretical investigation that is related to counting, Fourier Transforms, and Power Series. Guessing its exact shape is not something that comes very naturally on experimentalist intuition alone.

   In this week's lab, we will simply provide you with the right curve and continue from last week's lab on biased coins. Unless specified otherwise, you can assume the same configurations from last week's lab. In other words, the coin is biased with $P(\text{head}) = 0.7$, the number of tosses are ($k = 10, 100, 1000, 4000$), respectively, and the number of trials is $m = 1000$. Make sure you review the lab solution from Homework 10 before moving on.

   In addition, we will also look at the Birthday Paradox and Stirling's Approximation. Please come back to the last three parts of the lab when you are working on Question 8.

   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.

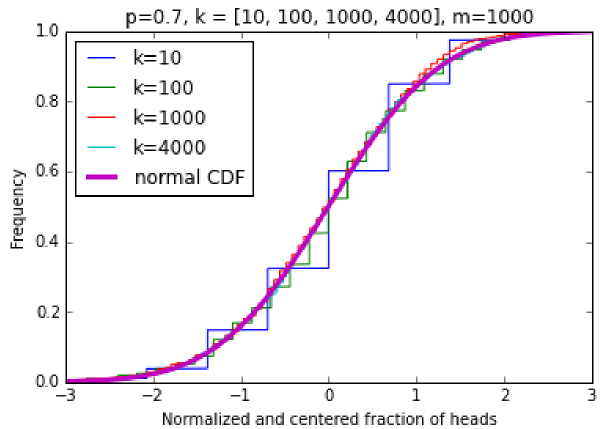
   Please download the IPython starter code from Piazza or the course webpage, and answer the following questions.

   (a) Plot $\int_{-\infty}^{d} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ overlaid with the normalized cliff-face shapes you had plotted in last week's lab. This integral is related to something called the Error Function. What do you observe?

   This is the heart of the Central Limit Theorem as applied to coin tosses.

   Hint: Implement the function `normal`, which takes a real number x and returns $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Then, implement the function `integrate_normal(d)`, which integrates the above function from $-\infty$ to d. In Python, you can use `scipy.integrate.quad`.

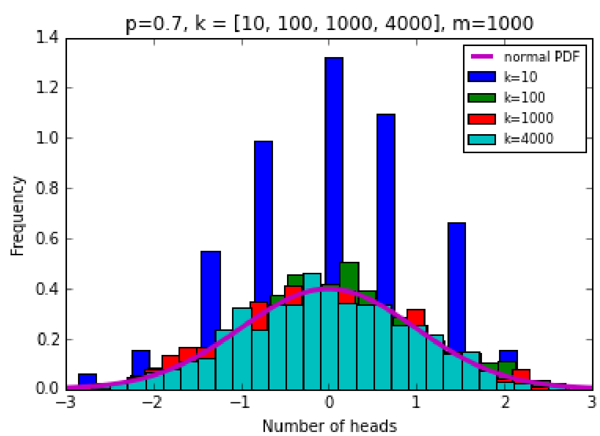
Solutions: We can see that the theoretical curve fits the shape perfectly, and the zigzags smooth out as k increases.



(b) Now, since you had realized earlier that the cliff-faces and the histograms have some natural relationship with each other, how would you naturally overlay a smooth plot of $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ to the normalized histograms. What does this mean?

Hint: There's a parameter for `plt.hist()` you learned in HW7 that you can use to normalize the histogram. Also, use a bin size of 0.2.

Solutions: For larger values of k, the bell-shaped curve overlays almost perfectly with the histograms. This suggests that integrating these histograms will yield the cliff-face shapes plotted above. It also suggests that the probability of getting certain numbers of heads is related to the bell-shaped curve.



(c) Another interesting pattern that you had seen in the previous Virtual Labs was the exponential drop in the frequencies of certain rare events. For an exponential drop, the most interesting thing is to understand the rate of the exponential — or the relevant slope on the Log-Linear plot.

For a coin with probability p of being heads, we are interested in the frequency by which tossing k such coins results in more than ak heads (where a is a number larger than p). We are interested in $p = 0.3, 0.7$ and $a = p+0.05$, $p+0.1$. Take $m = 1000$ and plot the natural log of the frequencies these deviations against k (ranging from 10 to 200). Approximately extract the slopes for all four of these.

Compare them in a table against the predictions of the following formula (which we will derive later in the course).

$$D(a||p) = a \ln \frac{a}{p} + (1-a) \ln \frac{1-a}{1-p}.$$

This expression is called the Kullback-Leibler divergence and is also called the relative entropy.

Finally, add $e^{-D(a||p)k}$ to the plots (there should be 4 of these) you have made as straight lines for immediate visual comparison. This straight line is called a "Chernoff Bound" on the probability in question.

What do you observe?

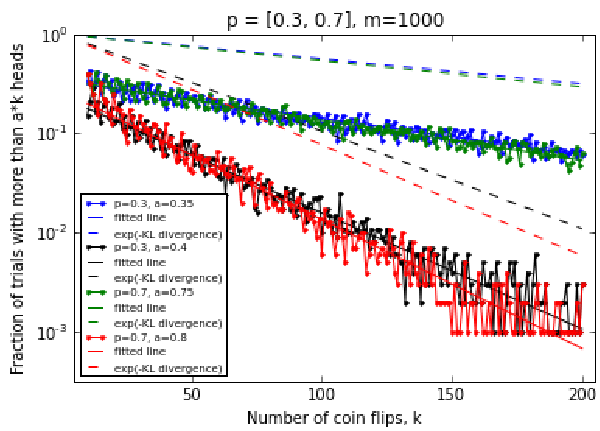
Hint: First, implement the function `KL`, which computes $D(a||p)$ using the given formula. To fit a line in Python, you can use `np.polyfit`.

There will probably be some 0 values, which will mess up this fitting, so you can replace the zeros with 10^{-3}.

Solutions: After fitting lines to the Log-Linear plots below, we obtain the following table.

p,a	fitted slope	$D(a\|\|p)$
0.3,0.35	-0.00865	0.00578
0.3,0.4	-0.02646	0.02258
0.7,0.75	-0.00922	0.00616
0.7,0.8	-0.02989	0.02573

We can see that the slope of the fitted lines are very close to the negative of the KL divergence. Also, the table values for $p+0.05$ are quite similar, and the values for $p+0.1$ are also similar. In the plot, we can see that the function $e^{-D(a||p)k}$ looks like it runs parallel to the fitted lines for each p,a pair.



(d) During your first week of Charm School (CS), you want to find fellow CS students who have the same birthday. Let's switch gears to an interesting problem studied in Lecture Note 12: the Birthday Paradox. This interesting phenomenon concerns the probability of two people in a group of m people having the same birthdays. This probability is given by

$$P(A^c) = 1 - \frac{365 \times 364 \times \ldots \times (365-m+1)}{365^m} = 1 - \frac{365!}{(365-m)!365^m}$$

where

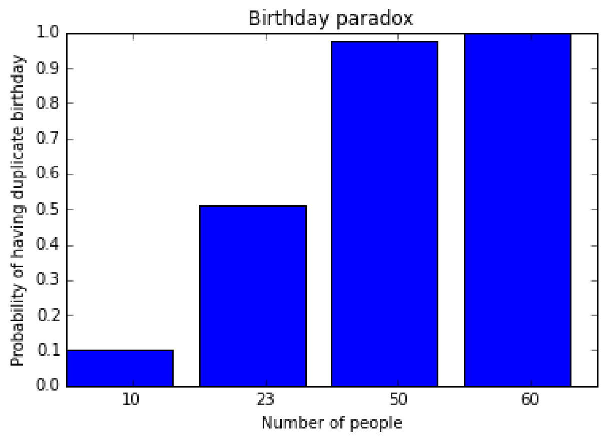
$$P(A) = \frac{365!}{(365-m)!365^m} = \left(1 - \frac{1}{365}\right) \times \left(1 - \frac{2}{365}\right) \times \ldots \times \left(1 - \frac{m-1}{365}\right)$$

and $P(A)$ is the probability that no two people have the same birthday.

For $m = 10, 23, 50, 60$, randomly generate birthdays by uniformly picking m numbers between 1 and 365. Do this 1000 times for each value of m. Record how many trials have at least 2 same birthdays. Plot this fraction vs. m using a bar chart. What do you observe?

Hint: First, implement the function `has_duplicate`, which returns True if a list contains any repeated element and False otherwise. Then, implement the function `gen_birthday(m)`, which generates random birthday for m people.

Solutions: As expected from Note 12, we see that for 23 people, the probability of at least two people having the same birthday is approximately 50%. For 60 people, it is approximately 99%!



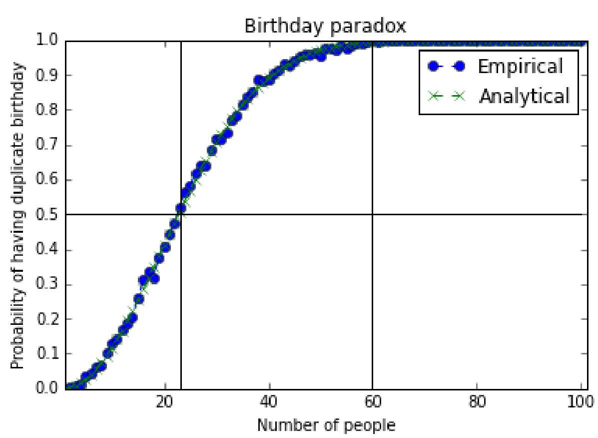
(e) We will now calculate the probability of having two people with the same birthday empirically, and plot the result against the expected probability, which is derived in Note 12.

Implement the function `birthday_formula(m)`, which calculates the probability of at least two people having the same birthday among m people.

Plot the empirical result (you can assume 1000 trials) v.s. the analytical result (`birthday_formula(m)`), for $m = [1, 100]$ people. What do you observe about the two curves? What happens at $m = 23$ and $m = 60$? Is this consistent with what we previously knew about the Birthday Paradox?

Solutions:

The two curves align with each other almost perfectly. At $m = 23$, we can see that there's a 50% probability of having at least two people with the same birthday. At $m = 60$, both curves slowly approach and converge to 1, so it's almost certain to have at least two people with the same birthday. This is consistent with what we previously studied in lecture – it is indeed a remarkable phenomenon!



(f) Now approximate $P(A)$ using Stirling's approximation for $n!$ and plot the approximated $P(A^c) = 1 - P(A)$ as a function of m. Stirling's approximation is given by

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Plot the analytical result from the previous part and the approximated result in the same figure. What do you observe?

Hint: Implement the function `birthday_stirling`, which computes the probability that no two people have the same birthday given that there are m birthdays using Stirling's approximation.

That said, don't use Stirling's approximation directly! Simplify your expression after using the approximation as much as possible before you implement the `birthday_stirling` function, or the large values will blow up your computer.

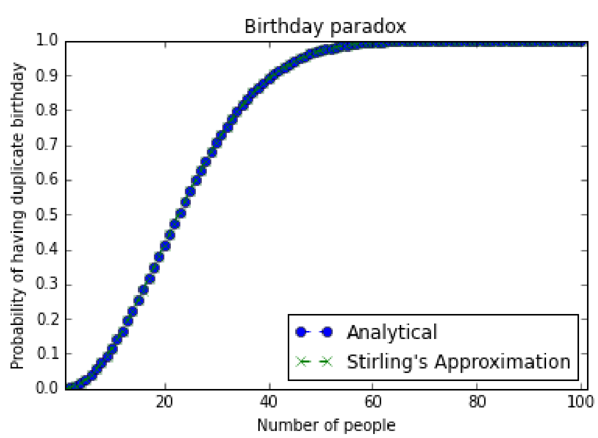
Solutions: First, let's derive the formula for Stirling's approximation to the Birthday Paradox. We know from the previous part that $P(A) = \frac{365!}{(365-m)!365^m}$. Let's apply Stirling's approximation to this expression.

$$\begin{aligned}
P(A) &= \frac{365!}{(365-m)!365^m} \\
&\approx \frac{\sqrt{2\pi(365)}(\frac{365}{e})^{365}}{\sqrt{2\pi(365-m)}(\frac{365-m}{e})^{365-m}365^m} \\
&= \sqrt{\frac{365}{365-m}}\frac{e^{365-m}}{e^{365}}\frac{365^{365-m}}{(365-m)^{365-m}} \\
&= \sqrt{\frac{365}{365-m}}e^{-m}\frac{(365-m)^{m-365}}{365^{m-365}} \\
&= \sqrt{\frac{365}{365-m}}e^{-m}(1-\frac{m}{365})^{m-365}
\end{aligned}$$

Plugging the final expression to the `birthday_stirling` function, we have

```
math.sqrt(365/(365-m))*math.pow(1-m/365, m-365)*math.pow(math.e,-m)
```

After plotting the two curves, we observe that they align perfectly on top of each other. This shows that Stirling's approximation is very powerful at approximating $n!$.



(g) Lastly, let's come back to the problem of counting the number of ways to throw m balls into n bins. Suppose the number of balls in each bin is a nonnegative integer, implement the function `permutation(m,n)`, which generates all possible permutations of throwing m balls into n bins. For example, `permutation(2,3)` should return

`[[0,0,2],[0,1,1],[0,2,0],[1,0,1],[1,1,0],[2,0,0]]`.

How would you change your implementation if we now require each bin to contain a positive number of balls?

Hint: Use recursion. You should have three base cases.

Solutions: See *lab11sol.pdf* for the implementation. If each bin must contain a positive number of balls, let's pretend to put one ball in each bin first. We now have $m-n$ balls left, and the problem is the same as before. In other words, we just need to set $m=m-n$ in the beginning and change nothing else in our implementation.

(h) Question 8, part (a)

(i) Question 8, part (h)

(j) Question 8, part (j)

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 `hw11.zip`.

3. Picking CS Classes

The EECS (Elegant Etiquette Charm School) department has d different classes being offered in Fall 2014. These include classes such as dressing etiquette, dining etiquette, and social etiquette, etc. Let's assume that all the classes are equally popular and each class has essentially unlimited seating! Suppose that c students are enrolled this semester and the registration system, EleBEARS (Elegant Bears), requires each student to choose a class s/he plans to attend.

(a) What is the probability that a given student chooses the first class, dressing etiquette?

Solutions: There are d different choices for each student. The probability of choosing "Dressing etiquette" class is $\frac{1}{d}$.

(b) What is the probability that a given class is chosen by no student?

Solutions: The probability a given student of choosing a class is $\frac{1}{d}$ and hence the probability of a given student not choosing the class is $\frac{d-1}{d}$. Each student chooses the class independently and there are c students. Hence the probability of no student choosing the class is $\left(\frac{d-1}{d}\right)^c$.

(c) If there are $d=20$ classes, what should c be in order for the probability to be at least one half that (at least) two students enroll in the same class?

Solutions: From Note 14, we know that for d bins, the probability of no collision is less than $\frac{1}{2}$ for approximately $\lfloor 1.177\sqrt{d} \rfloor$ balls. The problem of at least two students enrolling in the same class can be viewed as at least one collision ocurring for $d=20$ bins. Thus, in order for this probability to be at least $\frac{1}{2}$, the number of students needed is approximately $c = \lceil 1.177\sqrt{20} \rceil$, which gives us $c=5$ students.

If we list out the exact values of probabilities of no collision for varying values c, we get the following results:

For $c=1$, obviously $P(\text{no collision}) = 1$

For $c=2$, the second student has $20-1=19$ choices, so $P(\text{no collision}) = \frac{19}{20}$

For $c=3$, the second student has 19 choices and the third student has 18 choices, $P(\text{no collision}) = \frac{19\times 18}{20} = \frac{342}{400}$

Similarly,
For $c = 4$ $P(\text{no collision}) = \frac{19\times18\times17}{20^3} = \frac{5814}{8000}$
For $c = 5$, $P(\text{no collision}) = \frac{19\times18\times17\times16}{20^4} = \frac{93024}{160000}$
For $c = 6$, $P(\text{no collision}) = \frac{19\times18\times17\times16\times15}{20^5} = \frac{139536}{320000}$, which is finally less than $\frac{1}{2}$.
Hence it takes at least $c = 6$ students for the probability of collision to be at least half, very close to our approximation of $c = 5$ students.

4. Sock etiquette

In your second week of Charm School you learn that you should only wear matching pair of socks. In each pair, both socks must be of the same color and pattern. But all of them are in one big basket and now you have to take a pair out. Let's say you own n pairs of socks which are all perfectly distinguishable (no two pairs have the same color and pattern). You are now randomly picking one sock after the other without looking at which one you pick.

(a) How many distinct subsets of k socks are there?

Solutions: You could have interpreted the question to mean either that left and right socks are distinguishable from each other, or they are indistinguishable from each other.

For distinguishable left and right socks: We are picking k socks from $2n$ distinguishable socks, so there are $\binom{2n}{k} = \frac{2n!}{(2n-k)!k!}$ distinct subsets of k socks.

For indistinguishable left and right socks: This is a bit trickier. One way is to say that, for values $i \in \{0,\dots,\lfloor k/2\rfloor\}$, first choose i pairs and take both socks from these pairs, and then choose $k-2i$ out of the remaining $n-i$ pairs and take one sock from these pairs. This gives the answer

$$\sum_{i=0}^{\lfloor k/2\rfloor} \binom{n}{i}\binom{n-i}{k-2i}.$$

You could have also told a different story and come up with an equation that looks different, but still gives you the same answer. For example, for $i \in \{0,\dots,\lfloor k/2\rfloor\}$, you take $k-i$ distinct socks (no two from the same pair), and then choose i out of those $k-i$ and add their matching socks to the subset. This gives the answer

$$\sum_{i=0}^{\lfloor k/2\rfloor} \binom{n}{k-i}\binom{k-i}{i},$$

which is the same as the previous formula.

(b) How many distinct subsets of k socks which do not contain a pair are there?

Solutions:

When $k > n$, there are exactly 0 subsets of k distinct socks by the pigeonhole principle. Let's consider the case of interest when $k \leq n$.

For distinguishable left and right socks: One can construct a subset of $k \leq n$ socks which does not contain a pair by the following iterative process. Begin by picking any sock. While the number of picked socks is less than k, pick a sock belonging to a pair which has not yet appeared. In this process, we have $2n$ choices for the first sock, then $2(n-1)$ choices for the second one (as we can not pick the first sock again, nor pick the sock matching the first one), then $2(n-2)$ for the second sock, etc. Since we are counting subsets and the ordering does not matter, we divide everything by the number of ways to permute these k distinct socks, namely $k!$. Hence, there are

$$\begin{aligned}\frac{1}{k!}\times(2n\times 2(n-1)\times 2(n-2)\times\cdots\times 2(n-k+1)) &= \frac{1}{k!}\times\frac{2^k n!}{(n-k)!}\\ &= 2^k\binom{n}{k}\end{aligned}$$

such subsets.

For indistinguishable left and right socks: In this case, the answer is just $\binom{n}{k}$, since we can just take 1 sock from each pair and count the ways to make subsets of size k from this set.

(c) What is the probability of forming at least one pair when picking k socks out of the basket?

Solutions: For distinguishable left and right socks:

We have:

$$\begin{aligned}\mathbf{P}(\text{there exists one pair of socks in set of } k) &= 1-\mathbf{P}(\text{there is no pair of socks in the set of } k)\\ &= 1-\frac{|\{k \text{ (distinct) all unpaired socks}\}|}{|\{k \text{ (distinct) socks}\}|}\\ &= 1-\frac{2^k\binom{n}{k}}{\binom{2n}{k}}\\ &= 1-\frac{2^k n!(2n-k)!}{(n-k)!(2n)!}\end{aligned}$$

You will notice that the $k!$ terms simplify, so we could have counted the ordered versions of questions a) and b) instead and obtained the same result.

For indistinguishable left and right socks: In this case, we can't use a counting argument, since different combinations are no longer equally likely, e.g., the probability you choose 1 of sock type 1 and 1 of sock type 2 is not the same as the probability that you choose 2 of sock type 1. We can, however, still use an independence argument, just like we did for collisions of balls in bins (recall Discussion 11M). In this case, let a "collision" mean that we pick two socks from the same pair. We will pick socks one at a time.

The probability of having no collisions when we pick the first sock is 1. For the second sock, the probability of no collisions is $\frac{2n-2}{2n-1}$, since we have $2n-1$ socks left to choose from and $(2n-1)-1$ of them won't result in a collision. For the third sock we pick, the probability of no collisions is $\frac{2n-4}{2n-2}$, since we have $2n-2$ socks left to choose from and $(2n-2)-2$ of them won't result in a collision (we can't pick the two that we've already picked). Continuing this pattern, we can see that the probability of picking k socks with no collisions, assuming $k\le 2n$, is

$$\prod_{i=0}^{k-1}\frac{2n-2i}{2n-i} = 2^k\prod_{i=0}^{k-1}\frac{n-i}{2n-i} = 2^k\frac{n!/(n-k)!}{(2n)!/(2n-k)!} = \frac{2^k n!(2n-k)!}{(n-k)!(2n)!},$$

and therefore

$$\begin{aligned}\mathbf{P}(\text{there exists one pair of socks in set of } k) &= 1-\mathbf{P}(\text{there is no pair of socks in the set of } k)\\ &= 1-\frac{2^k n!(2n-k)!}{(n-k)!(2n)!},\end{aligned}$$

which is the same answer we got if we considered left and right socks to be distinguishable. It is important to recognize why, at an intuitive level, the probability is the same for both cases. We can pretend that someone secretly marks the right and left socks differently. But the person picking socks randomly can't see the marks and doesn't need to in order to pick a sock uniformly. A pair is a pair. So the probability must be the same.

(d) Now, in a different experiment, suppose there is exactly one sock of each pair in the basket (so there are n socks in the basket) and we sample (with replacement) k socks from the basket. What is the probability that we pick the same sock at least twice in the course of the experiment?

Solutions: This is basically the birthday problem with n days and k people. The number of ways to sample k socks with replacement is n^k. The number of ways to sample k socks with replacement without repetition is $n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!}$. Hence, the probability that we sample the same sock at least twice in the course of the experiment is:

$$1 - \frac{n!}{(n-k)!n^k}.$$

5. Drunk man

Imagine that you have a drunk man moving along the horizontal axis (that stretches from $x = -\infty$ to $x = +\infty$). At time $t = 0$, his position on this axis is $x = 0$. At each time point $t = 1$, $t = 2$, etc., the man moves forward (that is, $x(t+1) = x(t) + 1$) with probability 0.5, backward (that is, $x(t+1) = x(t) - 1$) with probability 0.3, and stays exactly where he is (that is, $x(t+1) = x(t)$) with probability 0.2.

(a) What are all his possible positions at time t, $t \geq 0$?

Solutions: Clearly, by time t, the man could have moved at most t positions to the right, and at most t positions to the left. Furthermore, within this range $[-t, t]$, the man could be occupying any integer position. Therefore, the possible values for the position $x(t)$ of the man at time t are exactly the integers in the closed range $[-t, t]$.

(b) Calculate the probability of each possible position at $t = 1$.

Solutions: Clearly, at time $t = 1$, the man could be either in position -1, or in position 0, or in position 1. We know the man starts at position 0 at $t = 0$, and at time $t = 1$, he has taken at most 1 step; if this step were taken backward (w.p. 0.3), he would be in position -1, and if this step were forward (w.p. 0.5), he would be in position $+1$. And if he had chosen to remain wherever he was (w.p. 0.2), he would be in position 0. There is no other way he could have been in any of these positions. So, his possible positions are $[-1, 0, 1]$, with probabilities $[0.3, 0.2, 0.5]$ respectively.

(c) Calculate the probability of each possible position at $t = 2$.

Solutions: From the discussion above, at time $t = 2$, the man can be in any one of the 5 positions $[-2, -1, 0, 1, 2]$. The probability associated with each of these positions can be calculated from the probabilities that we just computed above (for the man's position at time $t = 1$).

For example, what is the probability that the man is in position -2 at time 2? Clearly, this can happen under only one circumstance: the man should have been in position -1 at time 1, and moved backwards at time 2. Thus we have:

$$\begin{aligned}P(x(2)=-2) &= P\left(x(1)=-1 \cap \text{man moves backward at } t=2\right)\\ &= P\left(\text{man moves backward at } t=2 \mid x(1)=-1\right)\times P(x(1)=-1)\\ &= 0.3\times 0.3\\ &= 0.09.\end{aligned}$$

In general, using the law of total probability, we can write the probability that the man is in position i at time $t+1$ as

$$\begin{aligned}&P(x(t+1)=i)\\ =&P(x(t+1)=i|x(t)=i-1)\times P(x(t)=i-1)+P(x(t+1)=i|x(t)=i+1)\times P(x(t)=i+1)\\ &+P(x(t+1)=i|x(t)=i)\times P(x(t)=i)\\ =&0.5\times P(x(t)=i-1)+0.3\times P(x(t)=i+1)+0.2\times P(x(t)=i).\end{aligned} \tag{1}$$

Now, using (1),

$$\begin{aligned}&P(x(2)=-1)\\ =&P(x(2)=-1|x(1)=-2)\times P(x(1)=-2)+P(x(2)=-1|x(1)=0)\times P(x(1)=0)\\ &+P(x(2)=-1|x(1)=-1)\times P(x(1)=-1)\\ =&0+0.3\times 0.2+0.2\times 0.3\\ =&0.12.\end{aligned}$$

Doing this for the rest of the values, we get

$$\begin{aligned}&P(x(2)=0)\\ =&P(x(2)=0|x(1)=-1)\times P(x(1)=-1)+P(x(2)=0|x(1)=1)\times P(x(1)=1)\\ &+P(x(2)=0|x(1)=0)\times P(x(1)=0)\\ =&0.5\times 0.3+0.3\times 0.5+0.2\times 0.2\\ =&0.34.\end{aligned}$$

$$\begin{aligned}&P(x(2)=1)\\ =&P(x(2)=1|x(1)=0)\times P(x(1)=0)+P(x(2)=1|x(1)=2)\times P(x(1)=2)\\ &+P(x(2)=1|x(1)=1)\times P(x(1)=1)\\ =&0.5\times 0.2+0+0.2\times 0.5\\ =&0.2.\end{aligned}$$

$$\begin{aligned}&P(x(2)=2)\\ =&P(x(2)=2|x(1)=1)\times P(x(1)=1)+P(x(2)=2|x(1)=2)\times P(x(1)=2)\\ =&0.5\times 0.5+0\\ =&0.25.\end{aligned}$$

Notice that the 5 probabilities above add up to 1, as we would expect.

(d) Calculate the probability of each possible position at $t = 3$.

Solutions: From the discussion above, at time $t = 3$, the man can be in any one of the 7 positions -3, -2, -1, 0, 1, 2, or 3. The probability associated with each of these positions can be calculated from the probabilities that we just computed above (for the man's position at time $t = 2$).

The calculations are carried out in exactly the same way as in the previous part, by considering all possible ways in which the man can occupy position x at time 3, for each x satisfying $-3 \leq x \leq 3$.

$$\begin{aligned}
&P(x(3) = -3) \\
=&P(x(3) = -3|x(2) = -2) \times P(x(2) = -2) + P(x(3) = -3|x(2) = -3) \times P(x(2) = -3) \\
=&0.3 \times 0.09 + 0 \\
=&0.027.
\end{aligned}$$

$$\begin{aligned}
&P(x(3) = -2) \\
=&P(x(3) = -2|x(2) = -3) \times P(x(2) = -3) + P(x(3) = -2|x(2) = -1) \times P(x(2) = -1) \\
&+ P(x(3) = -2|x(2) = -2) \times P(x(2) = -2) \\
=&0.3 \times 0.12 + 0.2 \times 0.09 \\
=&0.054.
\end{aligned}$$

Proceeding in a similar fashion, the probabilities for the man to be in positions -3, -2, -1, 0, 1, 2, and 3 are 0.027, 0.054, 0.171, 0.188, 0.285, 0.15, and 0.125 respectively for $t = 3$. Again, as expected, these probabilities add up to 1.

(e) If you know the probability of each position at time t, how will you find the probabilities at time $t+1$?

Solutions: The solution to the previous part of the problem suggests a nice algorithm for computing the probability of each position the man can take at time $t+1$, provided these probabilities are known for time t.

Let X_t be the list of all possible positions that the man can be in at time t. From the arguments above, we know that:

$$X_t = [-t,\ -(t-1),\ \ldots,\ -1,\ 0,\ 1,\ \ldots,\ (t-1),\ t]\ .$$

Let Pt denote a list of probabilities corresponding to the positions X_t. Our goal is to find a way to calculate P_{t+1} (the *next probabilities*) given P_t (the *current probabilities*).

The figure above shows Python code for calculating the above next probabilities. The function `next_pvec` takes as input the current list of probabilities `pvec` (at time t), and values for `pf`, `pb`, and `pc` (the forward, backward, and "stay put" probabilities), and it produces as output a list of the next probabilities (at time $t+1$).

First of all, observe that the length of the list X_{t+1}, and hence P_{t+1} is two more than the length of X_t (hence P_t). This is because, at time $t+1$ the man can be in two additional possible positions that he could not have been in at time t.

Also, for each possible position at time $t+1$, there are at most 3 possible positions the man could have been in at time t. Therefore, the rules described above for multiplying the relevant probabilities and adding up these products generalize quite readily.

Thus, given the positional probabilities at time t, the man's positional probabilities at time $t+1$ can be readily calculated. And the man's initial position is known to be $x(0) = 0$. Therefore, starting from

```python
#! /usr/bin/env python2

import sys

def next_pvec (pvec, pf, pb, pc):

    qvec = []
    for idx in range(len(pvec)+2):
        q = pvec[idx]*pb if 0 <= idx < len(pvec) else 0
        q += pvec[idx-1]*pc if 0 <= idx-1 < len(pvec) else 0
        q += pvec[idx-2]*pf if 0 <= idx-2 < len(pvec) else 0
        qvec.append(q)

    return qvec


if __name__ == '__main__':

    (pf, pc, pb) = (0.5, 0.2, 0.3)
    tf = int(sys.argv[1])

    (t, pvec) = (0, [1])
    while t < tf:
        pvec = next_pvec (pvec, pf, pb, pc)
        t += 1

    print(pvec)
```

this initial condition, the positional probabilities can be calculated at any desired future time. Indeed, the `main` part of the above program does exactly this; it accepts a future time `tf` from the user and prints out a list of probabilities corresponding to every possible position the man can be in at time `tf`.

Note: Those of you who are familiar with Linear Algebra will readily recognize that the "next probabilities" list is simply a linear combination of the "current probabilities" list, which corresponds to pre-multiplying the current probabilities list by a (tall and thin) rectangular matrix. Indeed, this idea can be used to considerably speed-up the probability calculations above.

The Drunk Man has regained some control over his movement, and no longer stays in the same spot; he only moves forwards or backwards. More formally, let the Drunk Man's initial position be $x(0) = 0$. Every second, he either moves forward one pace or backwards one pace, *i.e.*, his position at time $t+1$ will be one of $x(t+1) = x(t)+1$ or $x(t+1) = x(t)-1$.

We want to compute the number of paths in which the Drunk Man returns to 0 at time t and it is his first return, *i.e.*, $x(t) = 0$ and $x(s) \neq 0$ for all s where $0 < s < t$. Note, we **no longer** care about probabilities. We are just counting paths here.

(f) How many paths can the Drunk Man take if he returns to 0 at $t = 6$ and it is his first return?

Solutions: We use an "F" to represent that the Drunk Man moves forward one pace and a "B" to represent that the Drunk Man moves backward one pace.

$\boxed{4}$ possible paths: FFFBBB, FFBFBB, BBBFFF, and BBFBFF. The last two paths can also be obtained by exchanging F's and B's in the first two paths.

(g) How many paths can the Drunk Man take if he returns to 0 at $t = 7$ and it is his first return?

Solutions: $\boxed{0}$ possible path because it needs the same number of forward paces and backward paces.

(h) How many paths can the Drunk Man take if he returns to 0 at $t = 8$ and it is his first return?

Solutions: $\boxed{10}$ possible paths: FFFFBBBB, FFFBFBBB, FFFBBFBB, FFBFFBBB, FFBFBFBB, and the other five by exchanging F's and B's.

(i) How many paths can the Drunk Man take if he returns to 0 at $t = 2n+1$ for $n \in \mathbb{N}$ and it is his first return?

Solutions: $\boxed{0}$ possible path because it needs the same number of forward paces and backward paces.

(j) How many paths can the Drunk Man take if he returns to 0 at $t = 2n+2$ for $n \in \mathbb{N}$ and it is his first return? (Hint: read http://en.wikipedia.org/wiki/Catalan_number and use any result there if you need.)

Solutions: From Wikipedia, "a Catalan number C_n is the number of monotonic paths along the edges of a grid with $n \times n$ square cells, which **do not pass above** the diagonal. A monotonic path is one which starts in the lower left corner, finishes in the upper right corner, and consists entirely of edges pointing rightwards or upwards." We can regard an F as an edge pointing rightwards, a B as an edge pointing upwards, and a possible path as a monotonic path of a grid with $(n+1) \times (n+1)$ square cells, which **does not touch** the diagonal. If the first pace is an F, then the $(2n+2)$-th path (the last pace) must be B; otherwise, the Drunk Man must have returned 0 before $t = 2n+2$. Therefore, in this case, we can focus on the second pace to the $(2n+1)$-th pace, and the number of possible paths is C_n because it is the number of monotonic paths, which do not pass above the diagonal of $n \times n$ square cells, *i.e.*, do not touch the diagonal of $(n+1) \times (n+1)$ square cells if the first pace is an F and the last pace is a B. Considering the other case that the first pace is a B and the last pace is an F, the total number of possible paths is $2C_n = \boxed{\frac{2}{n+1}\binom{2n}{n}}$.

6. An Identity on Integer Partitions

Let n be a positive integer. A partition of n is a way of writing n as a sum of positive integers. Partitions are considered equivalent under permutation of the summands, so that order of the summands does not matter. For example, 3 has exactly 3 partitions:

$$\begin{aligned} 3 &= 3 \\ &= 2+1 \\ &= 1+1+1 \end{aligned}$$

We will represent each partition p as a set of pairs (x,r) where the first element x represents a summand and the second element r is the number of times the summand appears, so that we have $n = \sum_{(x,r)\in p} rx$ for any partition p of n. We denote by $\mathscr{P}(n)$ the set of partitions of integer n. For example:

$$\mathscr{P}(3) = \{\{(3,1)\}, \{(2,1),(1,1)\}, \{(1,3)\}\}$$

In this problem, we will construct a combinatorial proof of the following identity:

$$\sum_{p\in\mathscr{P}(n)} \prod_{(x,r)\in p} \frac{1}{r!x^r} = 1$$

For example, for $n = 3$, the identity is saying:

$$\left(\frac{1}{1!3^1}\right) + \left(\frac{1}{1!2^1} \times \frac{1}{1!1^1}\right) + \left(\frac{1}{3!1^3}\right) = 1$$

(a) Make sure the above identity works for any $n \le 5$.

Solutions:

$$\mathscr{P}(1) = \{\{(1,1)\}\}$$
$$\sum_{p\in\mathscr{P}(1)} \prod_{(x,r)\in p} \frac{1}{r!x^r} = \left(\frac{1}{1!1^1}\right) = 1$$

$$\mathscr{P}(2) = \{\{(2,1)\},\{(1,2)\}\}$$
$$\sum_{p\in\mathscr{P}(2)} \prod_{(x,r)\in p} \frac{1}{r!x^r} = \left(\frac{1}{1!2^1}\right) + \left(\frac{1}{2!1^1}\right) = \frac{1}{2}+\frac{1}{2} = 1$$

$$\mathscr{P}(3) = \{\{(3,1)\},\{(2,1),(1,1)\},\{(1,3)\}\}$$
$$\sum_{p\in\mathscr{P}(3)} \prod_{(x,r)\in p} \frac{1}{r!x^r} = \left(\frac{1}{1!3^1}\right) + \left(\frac{1}{1!2^1}\times\frac{1}{1!1^1}\right) + \left(\frac{1}{3!1^3}\right) = \frac{1}{3}+\frac{1}{2}+\frac{1}{6} = \frac{2+3+1}{6} = 1$$

$$\mathscr{P}(4) = \{\{(4,1)\},\{(3,1),(1,1)\},\{(2,2)\},\{(2,1),(1,2)\},\{(1,4)\}\}$$
$$\sum_{p\in\mathscr{P}(4)} \prod_{(x,r)\in p} \frac{1}{r!x^r} = \left(\frac{1}{1!4^1}\right) + \left(\frac{1}{1!3^1}\times\frac{1}{1!1^1}\right) + \left(\frac{1}{2!2^2}\right) + \left(\frac{1}{1!2^1}\times\frac{1}{2!1^2}\right) + \left(\frac{1}{4!1^4}\right)$$
$$= \frac{1}{4}+\frac{1}{3}+\frac{1}{8}+\frac{1}{4}+\frac{1}{24} = \frac{6+8+3+6+1}{24} = 1$$

$$\mathscr{P}(5) = \{\{(5,1)\},\{(4,1),(1,1)\},\{(3,1),(2,1)\},\{(3,1),(1,2)\},\{(2,2),(1,1)\},\{(2,1),(1,3)\},\{(1,5)\}\}$$
$$\sum_{p\in\mathscr{P}(5)} \prod_{(x,r)\in p} \frac{1}{r!x^r} = \left(\frac{1}{1!5^1}\right) + \left(\frac{1}{1!4^1}\times\frac{1}{1!1^1}\right) + \left(\frac{1}{1!3^1}\times\frac{1}{1!2^1}\right) + \left(\frac{1}{1!3^1}\times\frac{1}{2!1^2}\right) + \left(\frac{1}{2!2^2}\times\frac{1}{1!1^1}\right)$$
$$+ \left(\frac{1}{1!2^1}\times\frac{1}{3!1^3}\right) + \left(\frac{1}{5!1^5}\right)$$
$$= \frac{1}{5}+\frac{1}{4}+\frac{1}{6}+\frac{1}{6}+\frac{1}{8}+\frac{1}{12}+\frac{1}{120} = \frac{24+30+20+20+15+10+1}{120} = 1$$

Let σ_n be the set of permutations over $\{1,2,\ldots,n\}$. Let $f \in \sigma_n$. We say that $(x_1\, x_2\, \ldots\, x_k)$ is a cycle of length k of f if and only if $f(x_1) = x_2, f(x_2) = x_3, \ldots, f(x_k) = x_1$. Note that $(x_1\, x_2\, \ldots\, x_k), (x_2\, x_3\, \ldots\, x_k\, x_1), (x_3\, x_4\, \ldots\, x_1\, x_2)$, $\ldots$ all represent the same cycle.

A familiar way to represent a permutation $f \in \sigma_n$ is to explicitly list the mapping $(x, f(x))$ for all $1 \le x \le n$. A different way to represent the same permutation is to list all its cycles. Consider Table 1 for an example of this.

x	1	2	3	4	5	6	7
$f(x)$	5	7	4	6	1	3	2

Table 1: A permutation $f \in \sigma_7$. The same permutation can also be represented by 1 cycle of length 3 and 2 cycles of length 2: (4 6 3), (2 7) and (1 5).

(b) Suppose we are working in σ_n. How many distinct cycles of length l can one construct?

Solutions: Here is a simple procedure to construct a length l cycle. Start by choosing the l elements which will be part of the cycle, then choose a permutation of these l elements and write down the ordered list of elements. This is one cycle. We shall however not forget to remove the l multiple counting of each cycle due to invariance by rotation, as we can essentially shift all the elements by $1 \leq i < l$ places around and obtain a different representation of the same cycle. Hence, the number we are looking for is:

$$\binom{n}{l} \times l! \times \frac{1}{l} = \frac{n!}{(n-l)!l} \tag{2}$$

(c) Let $(l_1, \dots, l_m)$, $(x_1, \dots, x_k)$, $(r_1, \dots, r_k)$ be 3 sequences of positive integers such that:

i. $\sum_{i=1}^{m} l_i = n$,

ii. $\sum_{j=1}^{k} x_j r_j = n$,

iii. For all $j \leq k$, there are exactly r_j distinct $i \leq m$ such that $l_i = x_j$.

How many distinct permutations in σ_n can be represented by a set of m cycles of length $l_1, \dots, l_m$? Express this number only in terms of n and the rs and xs. Be careful not to over count permutations.

Solutions: First, notice than if we are given m cycle lengths $l_1, \dots, l_m$ such that $\sum_{i=1}^{m} l_i = n$, we can indeed construct a permutation of n elements. Notice also that the x_is and r_is simply re-express the cycle lengths by grouping same cycle lengths (x_is) together and counting how many times each length occurs (r_is).

Again, let's give a procedure for constructing a permutation object out of our constraints and count the implied cardinality. The idea is to construct each cycle separately and reuse our result from the previous question. Start by constructing cycle 1 of length l_1. You have n elements to do so. Then construct cycle 2 of length l_2. Notice that you have only $n - l_1$ elements left to do so. Construct cycle 3 of length l_3. You have $n - (l_1 + l_2)$ left to do so. Keep constructing the cycles until the last one of length l_m with the remaining $n - (l_1 + l_2 + \cdots + l_{m-1})$ elements. This construction implies the following cardinality:

$$\frac{n!}{(n-l_1)!l_1} \times \frac{(n-l_1)!}{(n-l_1-l_2)!l_2} \times \cdots \times \frac{(n-(l_1+\cdots+l_{m-1}))!}{0!l_m}$$

We can greatly simplify this experssion by noticing the telescoping terms above and below the fraction bar. We get:

$$\frac{n!}{l_1 l_2 \dots l_m}$$

The expression above is a general form for sequence $l_1 l_2 \dots l_m$ where sum of the sequence is n. Now assume we have another sequence, ex: $x_1 x_1 x_1 x_2 x_2$ where sum of the sequence is n. Then again, we have:

$$\frac{n!}{(n-x_1)!x_1} \times \frac{(n-x_1)!}{(n-2x_1)!x_1} \times \frac{(n-2x_1)!}{(n-3x_1)!x_1} \times \frac{(n-3x_1)!}{(n-3x_1-x_2)!x_2} \times \frac{(n-3x_1-x_2)!}{(0!)!x_2}$$

Or

$$\frac{n!}{x_1^3 x_2^2}$$

Now we can extend it to r_1 x_1's, r_2 x_2's ... r_k x_k's. We can express it in equivalent terms of x_is and r_i:

$$\frac{n!}{x_1^{r_1}\dots x_k^{r_k}}$$

This is however not our final answer as we have over-counted all permutations by a (soon to be shown) constant factor. We have indeed pretended that the ordering of the cycles mattered by constructing cycles one by one starting from cycle 1 up to cycle m. However, the particular ordering of the cycles is meaningless, as (4 6 3) (2 7) (1 5) and (2 7) (4 6 3) (1 5) represent the same permutation from Table 1.

We now amend our construction according to this remark. We will enforce permutation distinctness by spilling out cycles in decreasing order of length size for example. We can do so by first sorting the l_is. While this fixes the problem of distinctness of permutations when all l_is are themselves distinct (as there is only one possible ordering and hence representation of a permutation in this case), it leaves the problem of the ordering of cycles of same length intact.

Removing these is however easy, as there are exactly $r_i!$ ways to permute r_i cycles of same size. Thus, we need to divide everything by $r_1! \times \dots r_k!$ to account for all equivalent representations of the same permutation object.

Hence, the cardinality we are looking for is:

$$\frac{1}{r_1!\dots r_k!} \times \frac{n!}{x_1^{r_1}\dots x_k^{r_k}} = n!\prod_{j=1}^{k}\frac{1}{r_j!x_j^{r_j}}$$

(d) You already know that $|\sigma_n| = n!$ by a simple counting argument. Now, use the previous question to count the elements of σ_n by using their cycle representation in order to prove the above identity.

Solutions: Decomposing permutations by the lengths of their cycles naturally introduces the integer partitions. Indeed, among all permutations of length n, we can consider the permutations that have a single cycle of length n, the permutations that have a single cycle of length $n-1$ and a cycle of length 1, etc. Using the result of the previous question, we have:

$$|\sigma_n| = \sum_{p\in\mathscr{P}(n)} n! \prod_{(x,r)\in p}\frac{1}{r!x^r}$$

And since $|\sigma_n| = n!$, simplifying by $n!$ gets us the above identity. QED.

7. Fibonacci Fashion

You have n accessories in your wardrobe, and you'd like to plan which ones to wear each day for the next t days. As a Charm School student, you know it isn't fashionable to wear the same accessories multiple days in a row. (Note that the same goes for clothing items in general). Therefore, you'd like to plan which accessories to wear each day represented by subsets $S_1, S_2, \dots, S_t$, where $S_1 \subseteq \{1,2,\dots,n\}$ and for $2 \le i \le t$, $S_i \subseteq \{1,2,\dots,n\}$ and S_i is disjoint from S_{i-1}.

(a) For $t \ge 1$, prove there are F_{t+2} binary strings of length t with no consecutive zeros (assume the Fibonacci sequence starts with $F_0 = 0$ and $F_1 = 1$).

Solutions: We will prove this by induction.

Base case: For $k = 1$, the only binary strings possible are 0 and 1. Therefore, there are two possible binary strings, and $F_{k+2} = F_3 = 2$. For $k = 2$, the binary strings possible are 11, 01, and 10, and we have $F_{k+2} = F_4 = 3$, so the identity holds.

Inductive hypothesis: Assume that for $k \geq 1$, there are F_{k+2} binary strings of length k with no consecutive zeros.

Inductive step: Consider the set of binary strings of length $k+1$ with no consecutive zeros. We can group these into two sets: those which end with 0, and those which end with 1.
For those that end with a 0, these can be constructed by taking the set of binary strings of length $k-1$ with no consecutive zeros and appending 10 to the end of them. Then by the inductive hypothesis, this set is of size F_{k+1}. For those that end with a 1, these can be constructed by taking the set of binary strings of length k with no consecutive zeros and appending a 1 to the end of them. Then by the inductive hypothesis, this set is of size F_{k+2}.
Since the union of these two subsets (those which end with 0 and those which end with 1) cover all possible elements in the set of binary strings of length $k+1$ with no consecutive zeros, the size of this set will be $F_{k+1} + F_{k+2} = F_{k+3}$. This thus proves the inductive hypothesis.

(b) Use a combinatorial proof to prove the following identity, which, for $t \geq 1$ and $n \geq 0$, gives the number of ways you can create subsets of your n accessories for the next t days such that no accessory is worn two days in a row:

$$\sum_{x_1 \geq 0} \sum_{x_2 \geq 0} \cdots \sum_{x_t \geq 0} \binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-x_2}{x_3}\cdots\binom{n-x_{t-1}}{x_t} = F_{t+2}^n.$$

Solutions: We first consider the left-hand-side of the identity. To create subsets of accessories that are consecutively disjoint with sizes $x_i = |S_i|$, $1 \leq i \leq n$, there are $\binom{n}{x_1}$ ways to create S_1, the subset of accessories you will wear on the first day. Then since S_2 must be disjoint from S_1, there are $\binom{n-x_1}{x_2}$ ways choose accessories to create S_2. Since S_3 must be disjoint from S_2, there are $\binom{n-x_2}{x_3}$ ways choose accessories to create S_3, and so on. Thus there are $\binom{n}{x_1}\binom{n-x_1}{x_2}\cdots\binom{n-x_{t-1}}{x_t}$ ways to create subsets of accessories $S_1, \ldots, S_t$ with respective sizes $x_1, \ldots, x_t$. Then altogether, $S_1, \ldots, S_t$ can be created in

$$\sum_{x_1 \geq 0} \sum_{x_2 \geq 0} \cdots \sum_{x_t \geq 0} \binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-x_2}{x_3}\cdots\binom{n-x_{t-1}}{x_t}$$

ways.
Now, consider the right-hand-side of the identity. Now for each accessory $i \in \{1, \ldots, n\}$, we will first decide which subsets $S_1, \ldots, S_t$ will contain accessory i, where we can't assign item i to consecutive subsets. For each accessory, we create a binary string of length t, where the leading digit represents S_1, the next digit represents S_2, and so on. We will say that a 0 in digit k means that we will wear the accessory on day k. Therefore, the number of ways we can assign accessory i to subsets $S_1, \ldots, S_t$ such that no two consecutive subsets both have accessory i is the same as the number of binary strings of length t with no consecutive zeros. Thus using the result in part (a), there are F_{t+2} ways to select the nonconsecutive subsets containing i among $S_1, \ldots, S_t$. Since we have n accessories, accessories $1, \ldots, n$ can be placed into subsets $S_1, \ldots, S_t$ in F_{t+2}^n ways.
This thus proves the identity.

8. Stirling's Approximation

In this question, suppose $n \in \mathbb{Z}^+$, we want to find approximations for $n!$. For the parts that are marked with **[VL]**, please complete your answer in the Virtual Lab skeleton. You can also use an online tool (*e.g.*, go to http://www.wolframalpha.com/ and type "plot $\ln x$") if you wish to.

(a) **[VL]** Plot the function $f(x) = \ln x$.

Solutions:

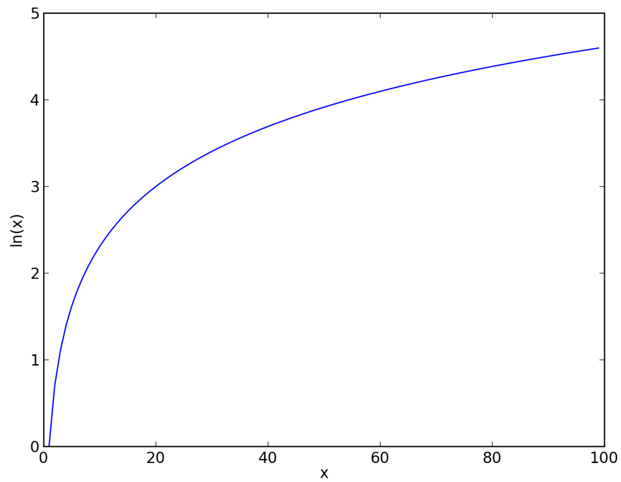


Figure 1: The plot of $\ln x$.

(b) For the following three questions, please note that $\ln x$ is strictly increasing and concave-$\cap$ because, when $x > 0$, its first and second derivatives are positive and negative, respectively. Concavity means that all line segments connecting two points on the function are below the function.
Suppose $n \in \mathbb{Z}^+$, use the plot to explain why

$$\ln 1 + \ln 2 + \ldots + \ln n \geq \int_1^n \ln x dx \tag{3}$$

Solutions:

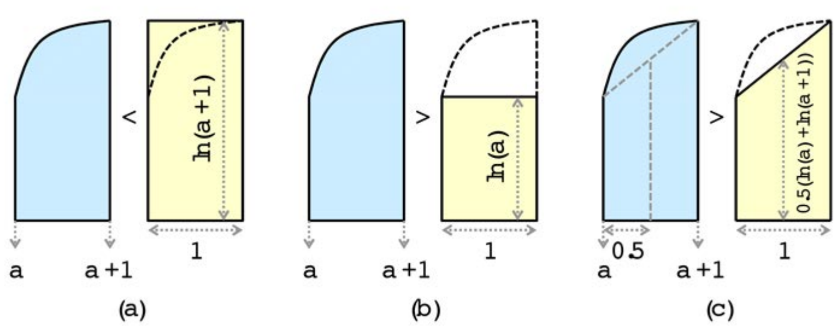


Figure 2: The area comparisons.

If $n = 1$, both sides are equal to 0. If $n \geq 2$, since $\ln x$ is strictly increasing and concave-$\cap$, we can see from Figure 2 (a) that the area of the left colored region is smaller than the right colored region, *i.e.*, $\int_a^{a+1} \ln x dx < \ln(a+1)$ where $a \geq 1$, so

$$\begin{aligned} \ln 1 + \ln 2 + \ldots + \ln n &= 0 + \ln 2 + \ln 3 + \ldots + \ln n \\ &> \int_1^2 \ln x dx + \int_2^3 \ln x dx + \ldots + \int_{n-1}^n \ln x dx \\ &= \int_1^n \ln x dx. \end{aligned}$$

(c) Suppose $n \in \mathbb{Z}^+$, use the plot to explain why

$$\ln 1 + \ln 2 + \ldots + \ln n < \int_1^{n+1} \ln x dx \quad (4)$$

Solutions:

Since $\ln x$ is strictly increasing and concave, we can see from Figure 2 (b) that the area of the left colored region is larger than the right colored region, *i.e.*, $\int_a^{a+1} \ln x dx > \ln a$ where $a \geq 1$, so

$$\begin{aligned} \ln 1 + \ln 2 + \ldots + \ln n &< \int_1^2 \ln x dx + \int_2^3 \ln x dx + \ldots + \int_n^{n+1} \ln x dx \\ &= \int_1^{n+1} \ln x dx. \end{aligned}$$

(d) Suppose $a \in \mathbb{Z}^+$, use the plot to explain why

$$\left(\frac{\ln a + \ln(a+1)}{2} \right) < \int_a^{a+1} \ln x dx \quad (5)$$

Solutions:

Since $\ln x$ is strictly increasing and concave, we can see from Figure 2 (c) that the area of the left colored region is larger than the right colored region, *i.e.*, $\int_a^{a+1} \ln x dx > \left(\frac{\ln a + \ln(a+1)}{2} \right)$ where $a \geq 1$.

(e) Use Equation ((3)) to prove $n! \geq e \left(\frac{n}{e}\right)^n$.

Solutions:

We have

$$\begin{aligned} \ln(n!) &= \ln 1 + \ln 2 + \ldots + \ln n \\ &\geq \int_1^n \ln x dx \\ &= (x \ln x - x)|_1^n \\ &= (n \ln n - n) - (0 - 1) \\ &= n \ln n - n + 1 \\ &= \ln(n^n) - \ln(e^n) + \ln e \\ &= \ln \left(e \left(\frac{n}{e} \right)^n \right), \end{aligned}$$

so $n! \geq e \left(\frac{n}{e}\right)^n$.

(f) Use Equation ((4)) to prove $n! \leq en \left(\frac{n}{e}\right)^n$ (Hint: If in this part you find yourself wishing you had $n-1!$ on the left-hand-side, try to prove an upper bound on $n-1!$ and use that to help you)

Solutions:

If $n=1$, $n! = en\left(\frac{n}{e}\right)^n = 1$. If $n>1$, we have

$$\begin{aligned}
\ln(n!) &= \ln 1 + \ln 2 + \ldots + \ln n \\
&= (\ln 1 + \ln 2 + \ldots + \ln(n-1)) + \ln n \\
&< \int_1^n \ln x dx + \ln n \\
&= (x\ln x - x)|_1^n + \ln n \\
&= (n\ln n - n) - (0-1) + \ln n \\
&= n\ln n - n + 1 + \ln n \\
&= \ln(n^n) - \ln(e^n) + \ln e + \ln n \\
&= \ln\left(en\left(\frac{n}{e}\right)^n\right),
\end{aligned}$$

so $n! < en\left(\frac{n}{e}\right)^n$ for $n>1$, and the claim $n! \le en\left(\frac{n}{e}\right)^n$ is proved.

(g) Use Equation ((5)) to prove $n! \le e\sqrt{n}\left(\frac{n}{e}\right)^n$, which is a tighter upper bound.

Solutions:

If $n=1$, $n! = e\sqrt{n}\left(\frac{n}{e}\right)^n = 1$. If $n>1$, we have

$$\begin{aligned}
\ln(n!) &= \ln 1 + \ln 2 + \ldots + \ln n \\
&= \frac{\ln 1}{2} + \left(\frac{\ln 1 + \ln 2}{2}\right) + \left(\frac{\ln 2 + \ln 3}{2}\right) + \ldots + \left(\frac{\ln(n-1) + \ln n}{2}\right) + \frac{\ln n}{2} \\
&< 0 + \int_1^2 \ln x dx + \int_2^3 \ln x dx + \ldots + \int_{n-1}^n \ln x dx + \frac{\ln n}{2} \\
&= \int_1^n \ln x dx + \frac{\ln n}{2} \\
&= (x\ln x - x)|_1^n + \frac{\ln n}{2} \\
&= (n\ln n - n) - (0-1) + \frac{\ln n}{2} \\
&= n\ln n - n + 1 + \frac{\ln n}{2} \\
&= \ln(n^n) - \ln(e^n) + \ln e + \ln(\sqrt{n}) \\
&= \ln\left(e\sqrt{n}\left(\frac{n}{e}\right)^n\right),
\end{aligned}$$

so $n! < e\sqrt{n}\left(\frac{n}{e}\right)^n$ for $n>1$, and the claim $n! \le e\sqrt{n}\left(\frac{n}{e}\right)^n$ is proved.

(h) **[VL]** The Stirling's approximation is usually written as $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ or a simpler version $n! \approx \left(\frac{n}{e}\right)^n$. Plot the function $f(n) = \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{n!}$. What do you observe?

Solutions:

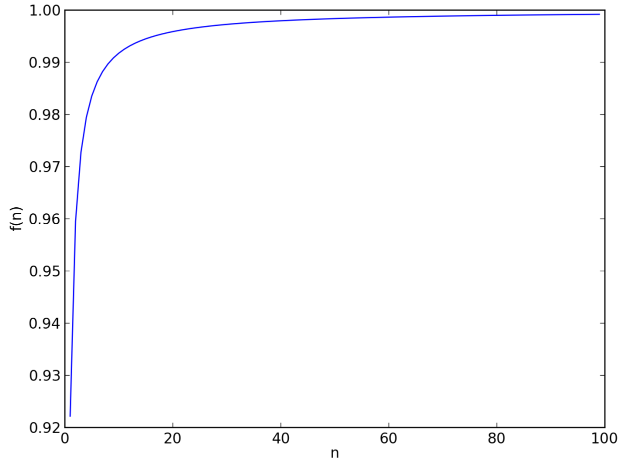


Figure 3: The plot of the ratio which is $f(n) = \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{n!}$.

The plot is shown in Figure 3. The function is closer to 1 as n increases.

(i) Suppose $m = \frac{k}{n}$, use m, n and apply the simpler version of the Stirling's approximation to rewrite $\binom{n}{k}$.

Solutions:

We have

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{(n-k)!k!} \\ &\approx \left(\frac{n^n}{e^n}\right)\left(\frac{e^{n-k}}{(n-k)^{n-k}}\right)\left(\frac{e^k}{k^k}\right) \\ &= \frac{n^n}{(n-k)^{n-k}k^k} \\ &= \left(\frac{n}{n-k}\right)^{n-k}\left(\frac{n}{k}\right)^k \\ &= \left(\frac{1}{1-m}\right)^{(1-m)n}\left(\frac{1}{m}\right)^{mn}. \end{aligned}$$

(j) **[VL]** Now, suppose $m_1 = \frac{k_1}{n} = 0.25$, $m_2 = \frac{k_2}{n} = 0.5$, and $m_3 = \frac{k_3}{n} = 0.75$, plot $\ln(\binom{n}{k_1})$, $\ln(\binom{n}{k_2})$, and $\ln(\binom{n}{k_3})$ as functions of n on a plot with linear-scaled axes. What do you observe?

Solutions:

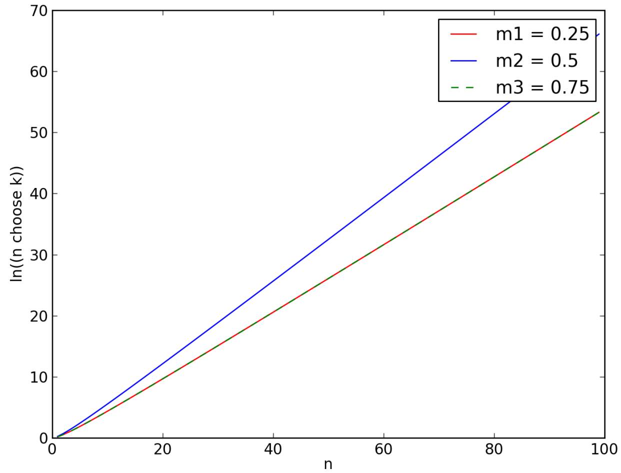


Figure 4: The plot of approximated $\ln(\binom{n}{k_i})$ where $m_1 = 0.25$, $m_2 = 0.5$, and $m_3 = 0.75$.

The plot is shown in Figure 4. The functions plotted are nearly linear! The functions with $m_1 = 0.25$ and $m_3 = 0.75$ overlap with each other.

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?