01204211 Discrete Mathematics
Lecture 16: Binomial Coefficients (3)

Jittat Fakcharoenphol

September 13, 2018

The binomial coefficients[1]

In this lecture, we discuss advanced counting with binomial coefficients. Then we shall study the function $\binom{n}{k}$ itself. First, let's see the actual value of the binomial coefficients $\binom{n}{k}$ for various values of n.

[1]This lecture mostly follows Chapter 3 of [LPV].

More on counting

We shall see more techniques for counting when we consider the following problems.

- How many anagrams does the word "KASETSARTUNIVERSITY" have? (They do not have to be real English words.)
- How can you give out n different presents to k students when student i has to get n_i pieces of presents?
- How many ways can you distribute n baht coins to k children?

Easy anagrams

- An anagram of a particular word is a word that uses the same set of alphabets. For example, the anagrams of ADD are ADD, DAD, and DDA.
- How many anagrams does "$ABCD$" have?
  - $4!$, because every permutation of A B C or D is a different anagram.

Harder anagrams

- How many anagrams does "$ABCC$" have? Is it $4!$?
  - This time we have to be careful because the answer of $4!$ is too large as it over counts many anagrams, i.e., it "distinguishes" the two C's.
  - Let's try to be concrete. How many times does "$CABC$" get counted in $4!$?
  - If we treat two C's differently as C_1 and C_2, we can see that $CABC$ is counted twice as C_1ABC_2 and C_2ABC_1. This is true for any anagram of $ABCC$.
  - Since each anagram is counted in $4!$ twice, the number of anagrams is $4!/2 = 4 \cdot 3 = 12$.

General anagrams

Let's try to use the same approach to count the anagram of $HELLOWORLD$. (It has 3 L's, 2 O's, H, E, W, R, and D.)

The number of permutation of alphabets in $HELLOWORLD$, treating each character differently is $10!$. However, each anagram is counted for $3!2!$ times because of the 3 copies of L and the 2 copies of O. Therefore, the number of anagrams is

$$\frac{10!}{3!2!}.$$

Distributing presents

I have 9 different presents. I want to give them to 3 students: A, B, and C. I want to give each student 3 presents. In how many ways can I do it?

- Let's think about the process of distributing the presents. We can first let A choose 3 presents, then B chooses the next 3 presents, and C chooses the last 3 presents. If we distinguish the order which each child chooses the presents, then there are $9!$ ways. However, in this case, we consider the distribution of presents, i.e., we consider the set of presents each child gets.
- To see how many times each distribution is counted in the $9!$ ways, we can let children form a line and let each child permute his or her presents. Each child has $3!$ choices. Thus, one distribution appears $3!3!3!$ times.
- Thus, the number of ways we can distribute presents is

$$\frac{9!}{3!3!3!}$$

Another way to look at the present distribution

- Let's look closely at a particular present distribution in the previous question. Let $\{1, 2, \ldots, 9\}$ be the set of presents.
- Consider the case where A gets $\{1, 3, 8\}$, B gets $\{2, 4, 6\}$, and C gets $\{5, 7, 9\}$.
- Another way to look at this distribution is to fix the order of the presents and see who gets each of the presents. Thus, the previous distribution is represented in the following table:

Presents	1	2	3	4	5	6	7	8	9
Children	A	B	A	B	C	B	C	A	C

- This is essentially an anagram problem. You can think of one particular way of present distribution as anagram of AAABBBCCC. Thus, we reach the same solution of

$$\frac{9!}{3!3!3!}.$$

Distributing identical presents

Now suppose that I have 9 identical presents. I want to give them to 3 students: A, B, and C. I want to give each student 3 presents. In how many ways can I do it?

- Note that when we state that the presents are identical, we mean that we do not distinguish them, i.e., the first present and the second present are indistinguishable.

Distributing coins (1)

I have 9 indentical coins. I want to give them to 3 students: A, B, and C. In how many ways can I do it so that each student gets at least one coin?

- Let's first try to organize the distribution of coins. We place all 9 coins in a line. We let the first student picks some coin, then the second student, then the last one.
- Since each coin is identical, we can let the first student picks the coin from the beginning of the line. Then the second one pick the next set of coins, and so on.
- One possible distribution is

$$\underbrace{oo}_{1}\underbrace{oooo}_{2}\underbrace{ooo}_{3}$$

- In how many ways can we do that?

Distributing coins (2)

The example below provides us with a hint on how to count.

$$\underbrace{oo}_{1}\underbrace{oooo}_{2}\underbrace{ooo}_{3}$$

Since all coins are identical, what matters are where the first student and the second student stop picking the coins. I.e, the previous example can be depicted as

$$oo|oooo|ooo$$

Thus, in how many ways can we do that?
Since there are 8 places we can mark starting points, and there are 2 starting points we have to place, then there are $\binom{8}{2}$ ways to do so.
This is a fairly surprising use of binomial coefficients.

Distributing coins (3)

Let's consider a general problem where we have n identical coins to give out to k students so that each student gets at least one coin. In how many ways can we do that?

Since there are $n-1$ places between n coins and we need to place $k-1$ starting points, there are $\binom{n-1}{k-1}$ ways to do so.

> There are $\binom{n-1}{k-1}$ ways to distribute n identical coins to k children so that each child get at least one coin.

Distributing coins (4)

I have 9 indentical coins. I want to give them to 3 students: A, B, and C. In how many ways can I do it, given that some student may not get any coins?

Identities in the Triangle

							1							
						1		1						
					1		2		1					
				1		3		3		1				
			1		4		6		4		1			
		1		5		10		10		5		1		
	1		6		15		20		15		6		1	
1		7		21		35		35		21		7		1

The binomial coefficients[2]

We now focus on the function $\binom{n}{k}$.
First, let's see the actual value of the binomial coefficients $\binom{n}{k}$ for various values of n.

[2]This lecture mostly follows Chapter 3 of [LPV].

What do you see?

- The function $\binom{n}{\cdot}$ is symmetric around $n/2$.
- Why? This is true because we know that $\binom{n}{k} = \binom{n}{n-k}$.
- The maximum is at the middle, i.e., when n is even the maximum is at $\binom{n}{n/2}$ and when n is odd, the maximum is at $\binom{n}{\lfloor n/2 \rfloor}$ and $\binom{n}{\lceil n/2 \rceil}$.
- Why? Can we prove that?

Largest in the middle

To understand the behavior of $\binom{n}{k}$ as k changes, let's look at two consecutive values:

$$\binom{n}{k} \heartsuit \binom{n}{k+1}$$

Let's write them out:

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \heartsuit \frac{n(n-1)(n-2)\cdots(n-k)}{(k+1)k!}.$$

Removing common terms, we can see that we are comparing these two terms:

$$1 \heartsuit \frac{n-k}{k+1} \Leftrightarrow k \heartsuit \frac{n-1}{2},$$

that is,

- if $k < (n-1)/2$, $\binom{n}{k} < \binom{n}{k+1}$; and
- if $k > (n-1)/2$, $\binom{n}{k} > \binom{n}{k+1}$.

How large is the middle $\binom{n}{n/2}$

Here, to simplify the calculation, we shall only consider the case when n is even. Let's try to estimate the value of $\binom{n}{n/2}$ by finding its upper and lower bounds.
A simple upper bound can be obtain using the fact that $\binom{n}{n/2}$ counts subsets of certain size:

$$\binom{n}{n/2} < 2^n.$$

We can also get a lower bound by noting that the maximum must be at least the average, i.e.,

$$\binom{n}{n/2} \geq \frac{2^n}{n+1}$$

Combining both bounds, we get that

$$\frac{2^n}{n+1} \leq \binom{n}{n/2} < 2^n.$$

Let's plug in $n = 200$, and calculate the number of digits to see how close these bounds.

$$27.80 \approx 200 \cdot \log 2 - \log 201 \leq \log \binom{n}{n/2} < 200 \cdot \log 2 \approx 30.10$$

Can we get a better approximation?
Yes, with Stirling's formula. (homework)

Concentration

- We know that the maximum of $\binom{n}{k}$ is obtained when $k = n/2$. From the graph, you can see that, as you move further from the middle, the value of the function drops rapidly.
- Since we consider even n, we let $2m = n$. One way to quantify how fast the values drop is to think about the ratio

$$\binom{2m}{m-t} \Big/ \binom{2m}{m}.$$

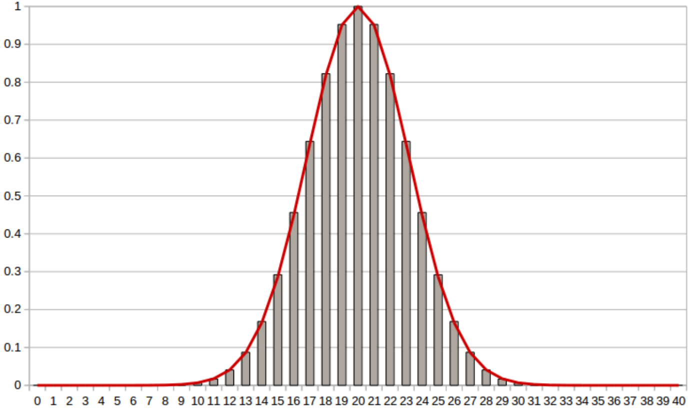
- In fact, it is known that

$$\binom{2m}{m-t} \Big/ \binom{2m}{m} \approx e^{-t^2/m}$$

- We will use our basic tools to obtain weaker bounds.

How close is the approximation?

The estimation $e^{-t^2/m}$ is extremely close as shown in the figure below, where the gray bars are the actual value of $\binom{2m}{m-t}/\binom{2m}{m}$ and the red line is $e^{-t^2/m}$.



The actual values

Because dealing with numbers less than 1 with logarithms is error-prone, we will work on the reciprocal. Let's try to calculate the ratio

$$\begin{aligned}\binom{2m}{m}\Big/\binom{2m}{m-t} &= \frac{(2m)!}{m!m!} \times \frac{(2m-m+t)!(m-t)!}{(2m)!} \\ &= \frac{(m+t)(m+t-1)\cdots(m+1)}{m(m-1)(m-2)\cdots(m-t+1)}.\end{aligned}$$

We can use the same logarithm trick. We have that the log of the ratio is

$$\ln\left(\frac{m+t}{m}\right) + \ln\left(\frac{m+t-1}{m-1}\right) + \cdots + \ln\left(\frac{m+1}{m-t+1}.\right).$$

Then we can apply the bounds we have for $\ln x$:

$$\frac{x-1}{x} \leq \ln x \leq x-1$$

The upper bound on the reciprocal

Each term in the sum is in this form $\ln((m-i)/(m+t-i))$. Applying the upper bound, we get

$$\ln\left(\frac{m+t-i}{m-i}\right) \leq \frac{m+t-i}{m-i} - 1 = \frac{m+t-i-m+i}{m-i} = \frac{t}{m-i}.$$

Let's sum them up to get

$$\ln\left(\frac{m+t}{m}\right) + \ln\left(\frac{m+t-1}{m-1}\right) + \cdots + \ln\left(\frac{m+1}{m-t+1}.\right)$$

$$\begin{aligned}
&\leq \quad \frac{t}{m} + \frac{t}{m-1} + \cdots + \frac{t}{m-t+1} \\
&\leq \quad \frac{t}{m-t+1} + \frac{t}{m-t+1} + \cdots + \frac{t}{m-t+1} \\
&= \quad \frac{t^2}{m-t+1}.
\end{aligned}$$

This implies that

$$\ln\left(\frac{(m+t)(m+t-1)\cdots(m+1)}{m(m-1)(m-2)\cdots(m-t+1)}\right) \le \frac{t^2}{m-t+1},$$

i.e.,

$$\begin{aligned}\binom{2m}{m}\Big/\binom{2m}{m-t} &= \left(\frac{(m+t)(m+t-1)\cdots(m+1)}{m(m-1)(m-2)\cdots(m-t+1)}\right) \\ &\le e^{t^2/(m-t+1)}.\end{aligned}$$

Taking the reciprocal, we get

$$e^{-t^2/(m-t+1)} \le \binom{2m}{m-t}\Big/\binom{2m}{m}.$$

Upper bounds

Using the same approach, we can show that

$$\binom{2m}{m-t} \Big/ \binom{2m}{m} \le e^{-t^2/(m+t)}.$$

Thus, we derived the estimates:

$$\boxed{e^{-t^2/(m-t+1)} \le \binom{2m}{m-t} \Big/ \binom{2m}{m} \le e^{-t^2/(m+t)},}$$

which is fairly close the the estimate of $e^{-t^2/m}$.

How fast?

- Let's return to the question on how fast do the values of the binomial coefficients decrease as you move further from the middle. Let's use the better estimate $\binom{2m}{m-t}/\binom{2m}{m} \approx e^{-t^2/m}$.
- Given a constant C, we want to estimate the value of t such that $\binom{2m}{m-t}$ is less than $\binom{2m}{m}/C$. (E.g., we can set $C = 2$ to see when the value drops by 50%.) Therefore, we want to find t such that

  $$1/C \geq \binom{2m}{m-t}/\binom{2m}{m} \approx e^{-t^2/m}$$

  Taking the logs, we get

  $$\ln 1/C = -\ln C \geq \ln \binom{2m}{m-t}/\binom{2m}{m} \approx -t^2/m.$$

  This is true when

  $$t \geq \sqrt{m \ln C}.$$

What does this means?

As an example, let $m = 20$ and $C = 2$. We know that when t is approximately $\sqrt{20 \cdot \ln 2} = 3.723$ the value of $\binom{2m}{m-t}$ drops by 50%.