NOTES ON SEPT 20 - 25, 2012

MIKE ZABROCKI

We considered rearranging letters of a word. I looked at the number of rearrangements of the word ANNOTATE. Consider rearrangements of the letters like TNTAAOEN or NEONATAT. I said that the following procedure will determine the word

- pick two positions from 8 for the letter A
- pick one position from the remaining 6 for the letter E
- pick two positions from the remaining 5 for the letter N
- pick one position from the remaining 3 for the letter O

the remaining two positions of the word will be filled with T's. That the set of rearrangements of the word ANNOTATE is in bijection with the sequences of subsets of $\{1, 2, \ldots, 8\}$ consisting of a subset of size 2, a subset of size 1, a subset of size 2 and a subset of size 1.

For example the word TNTAAOEN is sent under this bijection to $(\{4, 5\}, \{7\}, \{2, 8\}, \{6\})$. The number of such sequences is equal to

$$\binom{8}{2}\binom{6}{1}\binom{5}{2}\binom{3}{1} = \frac{8!}{2!6!}\frac{6!}{1!5!}\frac{5!}{2!3!}\frac{3!}{1!2!} = \frac{8!}{2!1!2!1!}$$

For this we define the notation we will call the multi-choose or multinomial coefficient. We will define $\binom{n}{k_1,k_2,\cdots,k_r}$ to be the number of ways of picking subsets of size $k_1, k_2 \ldots, k_r$ from an n element set For a sequence of integers $k_1, k_2, \ldots, k_r \geq 0$ such that $k_1 + k_2 + \cdots + k_r \leq n$, then

$$\begin{aligned}\binom{n}{k_1, k_2, \cdots, k_r} &= \binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3}\cdots\binom{n-k_1-k_2-\cdots-k_{r-1}}{k_r} \\ &= \frac{n!}{k_1!k_2!\cdots k_r!(n-k_1-k_2-\cdots-k_r)!}\ .\end{aligned}$$

If $k_1 + k_2 + \cdots + k_r > n$ then $\binom{n}{k_1,k_2,\cdots,k_r} = 0$.

There is another place where this coefficient arises. I assume that everyone is familiar with the binomial theorem which gives an expansion of $(1 + x)^n$ in terms of the binomial coefficients $\binom{n}{k}$. We have

$$(1+x)^n = \sum_{k\geq 0}\binom{n}{k}x^k = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n$$

for example, we have in particular

$$(1 + x)^4 = 1 + 4x + 6x^2 + 4x^3 + x^4 + 0x^5 + 0x^6 + 0x^7 + \ldots$$

The multinomial coefficient is a generalization of these coefficients. In fact, we have

$$(1 + x_1 + x_2 + \cdots + x_r)^n = \sum_{k_1+k_2+\cdots+k_r \leq n} \binom{n}{k_1, k_2, \ldots, k_r} x_1^{k_1} x_2^{k_2} \cdots x_r^{k_r}$$

With so many unknowns in this equation it is hard to appreciate this formula. But try an example. I can use the computer and see that $(1 + x + y)^4 =$

$$1+4\,x+4\,y+6\,x^2+12\,xy+6\,y^2+4\,x^3+12\,x^2y+12\,xy^2+4\,y^3+x^4+4\,x^3y+6\,x^2y^2+4\,xy^3+y^4$$

I can use this formula to see that $\binom{4}{1,2} = \frac{4!}{1!2!1!} = 12$ and I see that the coefficient of xy^2 in this expression is 12. If I want to answer a question like what is the coefficient of $x^7y^3z^9$ in the expression $(1+x+y+z)^{40}$ then I have a formula for this value, it is $\binom{40}{7,3,9} = \frac{40!}{7!3!9!21!}$ just as the binomial theorem tells me the coefficient of x^{19} in $(1 + x)^{40}$ is $\binom{40}{19} = \frac{40!}{19!21!}$.

I also wanted to know *how many* terms there are in the expression $(1+x_1+x_2+\cdots+x_r)^n$ (and not just an expression for the coefficient). In order to do this I looked at the number of monomials of degree d for $d = 0, 1, 2, 3, 4, 5, \ldots$ and when $r = 1, 2, 3, 4, \ldots$. Lets look at examples like

$$(1 + x)^5 = 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5$$
$$(1 + x)^6 = 1 + 6x + 15x^2 + 20x^3 + 15x^4 + 6x^5 + x^6$$

the number of terms of degree d is always 1 as long as d is less then or equal to the power n in $(1+x)^n$. Now lets try this for $r = 2$ variables. We computed $(1 + x + y)^4$, so lets look at

$$(1 + x + y)^5 = 1 + 5\,x + 5\,y + 10\,x^2 + 20\,xy + 10\,y^2 + 10\,x^3 + 30\,x^2y + 30\,xy^2 + 10\,y^3+$$
$$5\,x^4 + 20\,x^3y + 30\,x^2y^2 + 20\,xy^3 + 5\,y^4 + x^5 + 5\,x^4y + 10\,x^3y^2 + 10\,x^2y^3 + 5\,xy^4 + y^5$$

We see that at degree 0 there is 1 term, degree 1 there are two terms, degree 2 there are 3 terms, degree there are 4 terms, etc. (as long as $d \leq n$).

For three variables, the monomials of degree 1 are 3 monomials x, y, z; at degree 2 there are 6 monomials, $x^2, y^2, z^2, xy, xz, yz$; at degree 3 there are 10 monomials x^3, y^3, z^3, x^2y, x^2z, y^2x, y^2z, z^2x, z^2y, xyz; and one can conjecture that the pattern of the next value increases by 2,3,4,... continues.

If you write down a table, you see that the values we have computed thus far have a familiar pattern.

$r\backslash d$	0	1	2	3	4	5
1	1	1	1	1	1	1
2	1	2	3	4	5	6
3	1	3	6	10	15	21
4	1	4				

If you look closely at this table you see the binomial coefficients (Pascal's triangle). It might cause you to conjecture that the formula for the number of monomials in r variables of degree d is equal to $\binom{d+r-1}{d}$.

How do we explain this? First, notice that the exponents of a monomial can be translated to a list. $x^5y^3z^1w^6$ can be represented by $(5,3,1,6)$ without losing information. The degree of the monomial is equal to the sum of the entries in this list. Therefore we have shown:

$\#$ of monomials of degree d with r variables
$= \#$ of sequences of r non-negative integers that sum to d

Now there is another transformation that we can do in order to count these lists. Every sequence of non-negative integers can be translated into a sequence of dots $\bullet$ and bars $|$. The sequence $(a_1, a_2, \ldots, a_{r-1}, a_r)$ is in bijection with a_1 dots $\bullet$ followed by a bar $|$, a_2 dots $\bullet$ followed by a bar $|$, ..., a_{r-1} dots $\bullet$ followed by a bar $|$ followed by a_r dots $\bullet$. Notice that I don't need to finish with the bar because it would always be there so I leave it off. For example the sequence $(3,0,0,1,1,2,0)$ is sent to $\bullet\bullet\bullet|||\bullet|\bullet|\bullet\bullet|$ and we can recover the word consisting of d dots $\bullet$ and $r-1$ bars $|$ from the sequence of integers and we can recover the sequence of integers from the word, hence we have shown that these two things are in bijection. Therefore we have shown

$\#$ of monomials of degree d with r variables
$= \#$ of sequences of r non-negative integers that sum to d
$= \#$ of words with d symbols $\bullet$ and $r-1$ symbols $|$

This last set, we know how to count. In total there are $d+r-1$ letters in our word and d of them are $\bullet$ and $r-1$ of them are $|$, therefore we need only "choose" the subset of d positions where the $\bullet$s belong. Hence this is also in bijection with

$\#$ of monomials of degree d with r variables
$= \#$ of sequences of r non-negative integers that sum to d
$= \#$ of words with d symbols $\bullet$ and $r-1$ symbols $|$
$= \#$ subsets of size d of the integers $\{1,2,\ldots,d+r-1\}$

There is also one more set that this is in bijection with that I spent a while explaining. It is the number of ways of choosing d things from a set of size r where you are allowed to have repeat entries. Imagine we have an urn consisting of colored balls with r colors and you reach in and pull out d of them. A subset with repetition is just recording how many you got of each color. "I got 5 blue, 3 red, 1 green and 6 yellow" has the same information as a list of non-negative integers $(5,3,1,6)$. Hence a subset of the set $\{1,2,\ldots,r\}$ of size d where you allow repetitions is also in bijection with this collection of objects and we have:

$\#$ of monomials of degree d with r variables
$= \#$ of sequences of r non-negative integers that sum to d
$= \#$ of words with d symbols $\bullet$ and $r-1$ symbols $|$
$= \#$ subsets of size d of the integers $\{1, 2, \ldots, d+r-1\}$
$= \#$ multi-subsets (repetitions allowed) of size d of the integers $\{1, 2, \ldots, r\}$

I reviewed that we knew how to give a combinatorial interpretation to the binomial coefficient in a couple of odd ways.

Remark 1: How many non-negative integer solutions are there to the equation

$$x_1 + x_2 + \cdots + x_r = n?$$

Answer: $\binom{n+r-1}{n} = \binom{n+r-1}{r-1}$. Why? Think of a dots and bars argument and find a bijection from a solution to this equation represented as a sequence $(x_1, x_2, x_3, \ldots, x_r)$ and a sequence of n dots and $r-1$ bars.

Remark 2: How many paths are there in a lattice grid from $(0,0)$ to (n,m) with n steps $E = (1,0)$ and m steps $N = (0,1)$?
Answer: $\binom{n+m}{n} = \binom{n+m}{m}$. Why? Think of a lattice path in a grid with N and E steps and translate it into a word of letters N and E such that there are m letters N and n letters E. The number of such words is determined by the number of ways of choosing the positions of the E steps in the word.

Now prove

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

Proof 1: (using the combinatorial interpretation in terms of solutions to $x_1+x_2+\cdots+x_r = n$)
Let A be the set of non-negative integer solutions to the equation

$$x_1 + x_2 + \cdots + x_{n+1} = n$$

By Remark 1, we know that there are $\binom{2n}{n}$ such solutions. Let A_k be the subset of the solutions $(x_1, x_2, \ldots, x_n, x_{n+1})$ where $x_{n+1} = n-k$, then it must be that $x_1+x_2+\cdots+x_n = k$ and so also by Remark 1 we know that $|A_k| = \binom{n+k-1}{k}$ such solutions. Since

$0 \leq x_{n+1} \leq n$, we know that

$$A = \bigcup_{k=0}^{n} A_k$$

hence

$$\binom{2n}{n} = |A| = \sum_{k=0}^{n} |A_k| = \sum_{k=0}^{n} \binom{n+k-1}{k} \ .$$

Proof 2: (using the path combinatorial interpretation) Let A be the set of paths from $(0,0)$ to (n,n) using N and E steps. By Remark 2, there are $\binom{2n}{n}$ such paths. Let A_k be the set of paths in A such that the last horizontal step is at height k. A path in the set A_k consists of a path from $(0,0)$ to $(n-1,k)$ followed by a horizontal step to (n,k) followed by $n-k$ vertical steps. By Remark 2, there are $\binom{n+k-1}{k}$ paths from $(0,0)$ to $(n-1,k)$ and these paths determine completely the rest of the path, then $|A_k| = \binom{n+k-1}{k}$. Since $A = \bigcup_{k=0}^{n} A_k$,

$$\binom{2n}{n} = |A| = \sum_{k=0}^{n} |A_k| = \sum_{k=0}^{n} \binom{n+k-1}{k} \ .$$

I also found this example that I liked in the exercises in the book. Show

$$x^n - 1 = (x-1) + (x-1)x + (x-1)x^2 + \cdots + (x-1)x^{n-1} \ .$$

Proof: First start by assuming that $n, x \geq 1$ (otherwise the following argument won't make much sense). Let x represent the number of colors (one of which will be red), and n be a collection of n (ordered/distinct) rooms. Let A be the set of colorings of the n rooms with x colors such that all of the rooms are not red. There are x^n ways of coloring the rooms with no restriction on the colors since for each of the n rooms there is a choice of x colors. Therefor $|A| = x^n - 1$ since we exclude the possibility that all rooms are colored red. Now for $1 \leq k \leq n$, let A_k be the set of colorings of rooms in A such that the first room which is not red is the $n-k+1^{st}$ room (that is, A_1 is the set of rooms where all except the last room is red, A_2 is the set of colorings where the first $n-2$ rooms are red and the second to last is not red and the set A_n is the set of colorings of rooms where the first room is not red). We have broken up the set of colorings into n disjoint sets and $A = \bigcup_{k=1}^{n} A_k$. Moreover since A_k consists of $n-k-2$ red rooms, followed by a non-red room (in $x-1$ ways of coloring that room), followed by a coloring of the remaining $k-1$ rooms with x

possible colors, then $|A_k| = (x-1)x^{k-1}$ by the multiplication principle. Therefore

$$x^n - 1 = |A| = \sum_{k=1}^{n} |A_k| = \sum_{k=1}^{n} (x-1)x^{k-1} = (x-1) + (x-1)x + (x-1)x^2 + \cdots + (x-1)x^{n-1} .$$

Why make this proof so complicated? (we could have done it with a bit of algebra in 1/10th the time) What is worse that we have only managed to prove this identity for integers $x, n \geq 1$. It is hard to justify this sort of mathematical wank for this particular formula. It does give us a means of practicing when we are able to prove much more complicated combinatorial arguments. For example, to prove the Chu-Vandermonde formula

$$\binom{a+b+c+1}{a} = \sum_{k=0}^{a} \binom{k+b}{b} \binom{a-k+c}{c}$$

combinatorics is the easiest and it can even be used to make sure that your answer is right (see p. 50-51 in your book).

I should remark that even though we have given a proof of this identity for non-negative integers, we know that it is true for all integer values $n \geq 1$ and real numbers x. The reason is if we fix $n \geq 1$ then we have shown for all non-negative integer $x \geq 1$,

$$x^n - 1 - ((x-1) + (x-1)x + \cdots + (x-1)x^{n-1}) = 0.$$

But a polynomial of degree n has at most n zero points so if you tell me that $x^n - 1 - ((x-1) + (x-1)x + \cdots + (x-1)x^{n-1})$ is 0 for all positive integers, the only way that can happen is if it is the 0 polynomial and evaluates to 0 for all real values x.

For this argument I am using the following **FACT** which I have not shown you:
If $p(x)$ is a polynomial of degree $n > 0$, then there are at most n different real values x_i s.t. $p(x_i) = 0$.

I want to start in on sequences and so I gave a short introduction before we had to finish. Consider the following sequences

$$\begin{array}{ll}
\binom{0}{0}, \binom{1}{0}, \binom{2}{0}, \binom{3}{0}, \binom{4}{0}, \ldots & 1, 1, 1, 1, 1, 1, 1, \ldots \\
\binom{0}{1}, \binom{1}{1}, \binom{2}{1}, \binom{3}{1}, \binom{4}{1}, \ldots & 0, 1, 2, 3, 4, 5, 6 \ldots \\
\binom{0}{2}, \binom{1}{2}, \binom{2}{2}, \binom{3}{2}, \binom{4}{2}, \ldots & 0, 0, 1, 3, 6, 10, 15, \ldots \\
\binom{0}{3}, \binom{1}{3}, \binom{2}{3}, \binom{3}{3}, \binom{4}{3}, \ldots & 0, 0, 0, 1, 4, 10, 20, \ldots \\
\binom{0}{4}, \binom{1}{4}, \binom{2}{4}, \binom{3}{4}, \binom{4}{4}, \ldots & 0, 0, 0, 0, 1, 5, 15 \ldots
\end{array}$$

We have kind of written the table of binomial coefficients but we have selected the columns as our sequences to look at (we will also consider rows and diagonals too).

For any sequence of numbers $a_0, a_1, a_2, a_3, \ldots$ (if you have a finite sequence of numbers then put 0's at the end), we define the *generating function* of the sequence as the series

$$A(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \ .$$

Remark 1: The first mistake that a lot of people make about the generating function and the sequence is that they are not the same thing. For every sequence we have a generating function and for every generating function we can come up with a sequence. They are not equal. We use the phrases '...the generating function for/of a sequence...' and '...the sequence whose generating function is...' but please don't mix the two things up.

Remark 2: A generating function neither generates, nor (at least in our case) is it a function (although it looks like one). x is an indeterminate. x does not have a value. x is a placeholder. Sometimes I will use other variables instead of x, but those will also be unknowns and we are working in a space where it makes sense to manipulate the variable algebraically.

We all know the geometric series:

$$1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x}$$

(remember $(1-x)\cdot(1+x+x^2+x^3+x^4+\cdots) = 1+x+x^2+x^3+x^4+\cdots - x - x^2 - x^3 - x^4 - = 1$ so divide by $(1-x)$).

Notice that the second sequence has a generating function

$$x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + 6x^6 + \cdots$$

If you want a formula for it differentiate the first sequence to get $1+2x+3x^2+4x^3+5x^4+6x^5+\cdots$ and then multiply by x. That is,

$$x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + 6x^6 + \cdots = x\frac{d}{dx}\frac{1}{1-x} = \frac{x}{(1-x)^2}\ .$$