MATH 22

Lecture C: 9/9/2003

MORE COUNTING

When angry, count ten
before you speak;
If very angry, an hundred.
—Thomas Jefferson

When angry, count to four.
When very angry, swear.
—Mark Twain

Administrivia

- No office hours 9/11; office hours today instead, by request
- Homework / projects due today (homework in folders, projects in a separate pile)
- Questions about homework, project? Anyone want to see solutions?
- Questions/review from last lecture?
- http://denenberg.com/LectureC.pdf
- Response to: Post assignments in advance?
- Response to: Supplemental problems
- Comment: Projects are for exploration

Today: More about combinations, the Binomial Theorem & Pascal’s Triangle, other counting problems.

Examples of Combinations

How many distinct bridge hands are there?

Answer: C(52,13). 52 cards in a deck, choose any 13. (Doesn't work for poker!)

How many ways are there of choosing a baseball team from 18 people?

Answer: C(18,9).

How many ways are there of splitting 18 people into two baseball teams?

WRONG ANSWER: C(18,9). Choose 9 people for one of the teams; the other 9 form the second team.

Here's the trouble: At some point you select 1,2,3,...,9, so players 1 through 9 are on one team and 10 through 18 are on the other. But another selection consists of players 10,11,...,18, which results in the same two teams! See again Example 1.22 on p. 16, there done wrong.

The correct answer is C(18,9) / 2, since each way of splitting people into two teams is produced by *two* of the ways of choosing 9 out of the 18. This is yet another application of clumps: There are C(18,9) teams, but each way of splitting people into teams corresponds to a clump of two teams. The thing we want here is the number of clumps, which we get by dividing by 2.

Yet Another Example

How many ways are there of splitting 18 people into two baseball teams, the "Yankees" and the "Red Sox"?

Now the answer *is* C(18,9), since (presumably) putting players 1 through 9 on the Yankees and 10 through 18 on the Red Sox is *different* from the other way around!

How many ways are there of choosing the three winners of a horserace if there are ten horses running?

The answer is *not* C(10,3), which is the number of ways of picking three horses *where order doesn't matter*. But in a horserace, order *does* matter; A to win, B to show, C to place is much different than C to win, A to show, and B to place. So we have to go back to permutations. The correct answer is P(10,3).

Comment: Generally, permutations are the tool of choice when order is important (horserace, words) and combinations when order is unimportant (poker hands, lottery tickets). But you must not blindly apply this rule! Often it's not so obvious whether to use permutations or combinations, and sometimes we need both. You must understand the problem and the tools!

Combinatorial Identities

From the end of last lecture:

$$C(n,n) = 1$$

(There's only one way to choose all the members of a set of n.)

$$C(n,0) = 1$$

(There's only one way to choose zero members of a set of n.)

$$C(n,r) = C(n,n-r)$$

(To choose r members from a set of n, you can equally well choose $n-r$ members to leave out.)

Comment: It's convenient (and intuitively correct) to set $C(n,r) = 0$ whenever $r < 0$ or $r > n$.

Another Identity

$$C(n+1,r) = C(n,r) + C(n,r-1)$$

Useless, unenlightening proof:

Massage $n!/r!(n-r)! + n!/(r-1)!(n-r+1)!$
algebraically until it turns into $(n+1)!/r!(n+1-r)!$

(But you should be sure you can do this; it's good practice in handling factorials.)

A proof that promotes understanding:

$C(n+1,r)$ is the number of ways of choosing r items out of a set of $n+1$, order unimportant and without repetitions.

Call one of the $n+1$ items X. There are two ways we can choose the r items:

Ignore X; pick r items from among the other n items. This can be done in $C(n,r)$ ways.

Choose X plus $r-1$ of the other n items. This can be done in $C(n,r-1)$ ways.

These two methods don't overlap, so we can add the results together. Done!

And Another

$C(n,0) + C(n,1) + C(n,2) + \ldots + C(n,n) = 2^n$

We prove this identity by showing that each side is the number of subsets of a set of n elements (are we sure we know what this means?).

LHS: The number of subsets of a set of size n is equal to the number of subsets with size 0, plus the number of subsets with size 1, plus the number of subsets with size 2, and so forth, up to the number of subsets of size n.
But the number of subsets with size k is just the number of ways of choosing k elements from a set of n (order unimportant, no repetitions), which is $C(n,k)$. So the LHS above is the total number of subsets of a set of size n.

RHS: To make a subset of a set of size n, we consider the elements one after the other and answer the question "in or out"? We need to make n independent decisions, and for each there are 2 possible choices; each way of doing this leads to a different subset. By the Rule of Product, the number of ways of doing it is 2^n.

(This is an awesome proof technique: Show that two expressions are equal by showing that they each count the same thing!)

The Binomial Theorem

As we (should) know,

$$(x+y)^2 = x^2 + 2xy + y^2$$

Let's recall why. We get this by "cross multiplying" $(x+y)(x+y)$, giving four terms: xx, xy, yx, and yy. Of course xy and yx are the same (an x times a y) so they can be combined into $2xy$. (What is a term?)

Now what about

$$(x+y)^n = (x+y)(x+y)(x+y)\ldots(x+y)$$

The answer is going to be the sum of cross-multiplying all possible terms. That is: Pick one thing, either an x or a y, from each pair. Multiply these n things together; the answer is $x^k y^j$ for some k and j. Repeat this procedure for all possible ways of picking, and add all of the results together!

Let's look at one term $x^k y^j$. Obviously $0 \le j,k \le n$ (are we sure what this means?) and also $j+k = n$. So we might as well write a term as $x^k y^{n-k}$. Also, similar terms (those with the same k) can be collected. So the answer is

$$A_0 x^0 y^n + A_1 x^1 y^{n-1} + A_2 x^2 y^{n-2} + \ldots + A_{n-1} x^{n-1} y^1 + A_n x^n y^0$$

and all we need to do is calculate the coefficients A_k. (Pause to understand, and to see the values of A_0 and A_n.)

Binomial Thm, cont.

How many cross-multiplications yield $x^k y^{n-k}$ for given k?

Answer: Think of the n copies of $(x+y)$ as n distinct items. Now choose k of those items to be the ones donating y to the term. The remaining $n–k$ items donate an x, and the result is $x^k y^{n-k}$.

How many ways are there to choose the k items to donate y? Obviously there are $C(n,k)$ such ways: Order doesn't matter, and each item can be chosen at most once. So the coefficient of $x^k y^{n-k}$ is $C(n,k)$.

We have proven the Binomial Theorem:

$$(x+y)^n = C(n,0)x^0y^n + C(n,1)x^1y^{n-1} + C(n,2)x^2y^{n-2} +$$
$$+ \ldots + C(n,k)\, x^ky^{n-k} + \ldots + C(n,n\text{-}1)x^{n-1}y^1 + C(n,n)x^ny^0$$

You should verify that $(x+y)^2 = x^2+2xy+y^2$ is the special case of this when n=2.

Three Comments

Comment 1: Take the Binomial Theorem,

$(x+y)^n = C(n,0)x^0y^n + C(n,1)x^1y^{n-1} + \ldots + C(n,n)x^ny^0$

and substitute $x = y = 1$. The result is

$$2^n = C(n,0) + C(n,1) + \ldots + C(n,n).$$

Look familiar?

Comment 2: Now substitute $x = -1$, $y = 1$. Result:

$$0^n = 0 = C(n,0) - C(n,1) + C(n,2) - C(n,3) + \ldots$$

$$C(n,0)+C(n,2)+C(n,4)+\ldots = C(n,1)+C(n,3)+C(n,5)+\ldots$$

Can you prove this with a counting argument?

Hint: If n is odd the argument is totally trivial.

Comment 3: What if we're interested in

$$(x + y + z)^n$$

or more generally in

$$(x_1 + x_2 + x_3 + \ldots + x_k)^n$$

Same idea: We get, via complete cross-multiplication, terms of the form $x^py^qz^r$ where $0 \le p,q,r \le n$ and $p+q+r = n$, and the only problem is to count the number of terms for each triple of exponents. The result is called the Multinomial Theorem: the coefficient of $x^py^qz^r$ is $n!/p!q!r!$ (full derivation is in the book).

Pascal's Triangle

$$
\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
\cdots
\end{array}
$$

The r^{th} entry in the n^{th} row of Pascal's Triangle is $C(n,r)$.
(The top row is the zeroth row, containing only $C(0,0)$.)
Each entry is the sum of the two above it, according to

$$C(n,r) = C(n\text{-}1,r\text{-}1) + C(n\text{-}1,\mathrm{r})$$

which we saw earlier in slightly different form.
Each row is a sequence of binomial coefficients, e.g.

$$(x+y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$$

$$(x+y)^1 = x + y$$

$$(x+y)^0 = 1$$

Pascal's Triangle has lots of pleasant properties, some of which you will be exploring in a project (I think).

Combinations *with repetition*

How many ways are there to select r objects from a set of n items, where order is unimportant but *repetitions* are permitted?

Example: How many ways can I pick a dozen bagels, if the bagel shop makes twenty kinds?

If all the bagels have to be different, the answer is easy: C(20,12). But we're not just choosing twelve items from a set of twenty, because we can pick six egg bagels, three plain, etc.

Note that here it makes sense to have $r>n$, which doesn't make sense when repetitions aren't permitted. That is, $C(n,r)$ is always 0 if $n<r$ --- how can you choose 10 objects from a set of 5? But with repetitions it can be done: I *can* buy 10 bagels from a shop that makes 5 kinds of bagels!

(Order is still unimportant here. I can pick two plain then two egg, or instead a plain, then an egg, then a plain, then an egg. All that matters is what bagels are in my bag at the end!)

How shall we proceed?

Combs w/reps, cont.

Method I: Make a little "picture" that represents a choice. A picture consists of:

- Zero to r circles, denoting bagels of the first flavor, then
- A comma to separate flavor 1 from flavor 2, then
- Zero to r circles denoting bagels of flavor 2, then
- A comma to separate flavor 2 from flavor 3,

. . . and so forth up, up to . . .

- Zero to r circles denoting bagels of flavor n.

Here are some pictures for n=5 and r = 10 (that is, picking ten bagels from five flavors):

O O O , O O , O , O O O , O

(3 of the first flavor, 2 of the second, . . ., 1 of the fifth)

O O , , O O O , O O O O O , O

(2 of the first flavor, none of the second, etc.)

, , , O O O O O O O O O O ,

(all ten of the fourth flavor, obviously garlic)

Now we're just going to count the number of pictures!

Method I, cont.

Here are the critical points about the pictures:

• Each picture has exactly r circles (10 in the example); this represents the number of bagels to be chosen.

• Each picture has exactly n–1 commas (4 in the example); this is *one less than* the number of flavors. It's one less because the commas go *between* flavors. We'll see in a second why we want this.

• Each way of choosing the bagels leads to a single, unique picture. (Pretty obvious.)

• Any way at all of slapping down r circles and n–1 commas corresponds to exactly one way of choosing bagels! (Not necessarily so obvious, but true. Wouldn't be true if we put commas on the outside.)

Conclusion: If we count the number of ways of writing down r circles and n–1 commas, we have counted the number of ways of picking r bagels from n flavors!

And how many such ways are there? Each picture has $r+n-1$ marks where each mark is either a circle or a comma. The number of pictures is just the number of ways of picking which r of these marks are the circles!

So the answer, the number of ways of picking r items from a field of n with repetitions, is $C(n+r-1,r)$.

C with R, Method II

A completely different way of solving the same problem:

Label the flavors as numbers from 1 to n, that is, 1 represents the first flavor, 2 the second, and so forth.

Represent a choice of r bagels as an ordered *sequence* of numbers; each number is a bagel type (from 1 to n) and the sequence has length r. So, for example,

$$<3, 1, 4, 2, 5, 5, 2, 1, 1, 2>$$

represents choosing a bagel of type 3, then one of type 1, then one of type 4, etc.

Problem: This representation incorrectly makes significant the *order* in which we choose bagels! For example,

$$<5, 2, 1, 3, 4, 5, 1, 2, 2, 1>$$

is a different sequence, but encodes the same choice of bagels. We can repair this by always representing bagel choices as *sorted* sequences. That is, we wouldn't use either sequence above, but rather

$$<1, 1, 1, 2, 2, 2, 3, 4, 5, 5>$$

to represent this choice of bagels. The question now is to count the number of sorted sequences of length r, each entry of which is an integer from 1 to n.

Method II, cont.

Here's the trick: Given any sequence, let's make something which we'll call the *fixed* sequence ("fixed" in the sense of "repaired".) We "fix" a sequence by adding

$$<0,1,2,3,4, \ldots, r-1>$$

to it. Fixing the example from the previous slide gives

$$<1, 2, 3, 5, 6, 7, 9, 11, 13, 14>$$

Two things are true about any fixed sequence:

- Its elements are all between 1 and $n+r-1$.
- It's always *increasing* (since it's the sum of an increasing and a nondecreasing sequence).

We know that any choice of bagels corresponds to a unique fixed sequence. Now we assert that each fixed sequence corresponds to a unique choice of bagels!

How do we know this? We can *unfix* any fixed sequence by subtracting $<0,1,\ldots,r-1>$ from it (this can never go negative) and get back to some choice of bagels.

So we can solve the original problem by counting the fixed sequences. And a fixed sequence is just any choice of n numbers between 1 and $n+r-1$, order unimportant (since we sort it anyway) and *without* repetitions (since fixed sequences never have duplicates).

And, of course, this is just $C(n+r-1, r)$, as it should be.

Example: Backgammon

We once said that there are 36 ways that two dice can fall, but that a backgammon player doesn't care about 36 different possibilities (because it doesn't matter whether the dice fall 3&5 or 5&3, for example).

So how many ways can two dice fall in backgammon?

To get an answer, we rephrase the question like this: We have n=6 objects (namely, the possible ways a die can fall) and we want to select r=2 things from this set (namely, the value on each die) *with repetitions permitted* (because the dice can have the same value).

By our result, the answer is

$$C(6+2-1, 2) = C(7,2) = 21.$$

More generally, in how many different ways can k identical dice fall? Answer: Now we're asking to choose r=k things from a field of n=6, with repetitions allowed but with order unimportant, so the answer is

$$C(6+k-1, k)$$

A Variation

How many ways are there of buying r bagels from a bakery that offers n kinds of bagels, if we insist on buying *at least one of each kind*?

This is pretty simple. Start by picking one bagel of each kind—there's only one way to do this, and you have to do it anyway. Now we need to count the number of ways of picking $r–n$ bagels from a set of n kinds of bagels, order unimportant and repetitions permitted. By previous work, the answer is C(r–1, r–n).

Notice that when n=r or n=1 the answer is always 1, and when r<n the answer is always 0. Is this right?

Generalization: How many ways are there of buying r bagels from a bakery that offers n kinds of bagels, if we insist on buying *at least* q *bagels of each kind*? Clearly the answer is 0 when r<qn.

You should be able to finish this one off yourself. (And if it were up to me, you'd have to.)

Summary

How many ways are there to select r things from a set of n distinct items, . . .

• with order important, and repetitions permitted?
Answer: n^r

• with order important, and no repetitions permitted?
Answer: $P(n,r)$

• with order unimportant, and no repetitions permitted?
Answer: $C(n,r)$

• with order unimportant, and repetitions permitted?
Answer: $C(n+r-1, r)$

• with order unimportant, repetitions permitted, and at least 1 of each of the n selected?

Answer: $C(r-1, r-n)$

If the n items are not distinct, but there are m_1 of one kind, m_2 of a second kind, . . . , m_k of the k^{th} kind, where the sum of the m_k is n, then the number of ways of permuting all n items is $n! \,/\, m_1!m_2!...m_k!$ (and we don't know how to count anything else). [MISSISSIPPI]

NB: "repetitions permitted" is also meant when we say "with replacement"; the idea is that you choose a thing, but then put that thing back into the pot so that it can be chosen again.

Urn and Ball Problems

Urn and ball problems have been the bane of generations of students of combinatorics. We're looking at them here just as an exercise in distinguishability.

How many ways are there of putting r balls into n urns? The answer depends on whether you can tell the balls and the urns apart!

Suppose that both balls and urns are *distinguishable*. We simply need to choose one of the n urns for each ball, i.e., a choice of n, then another choice of n, for each of the r balls. By rule of product, the answer is n^r.

(What happens if we permit at most one ball per bin?)

Now suppose that the urns are distinguishable, but the balls are identical. This is the same as choosing r things from a set of n, with repetitions permitted. (We must choose a set of r urns for the balls to go into.) So the answer is $C(n+r-1, n)$.

(Again, what's the answer with at most one ball per bin?)

Finally, if the urns are *indistinguishable*, we don't yet have the techniques to solve the problem whether or not the balls are distinguishable! Wait and see.

Application:
Solutions of Equations

How many solutions are there to the equation

$$x_1 + x_2 + x_3 = 8$$

where each of the x_i is a nonnegative integer?

Write a solution as (x_1, x_2, x_3); some examples of solutions are (4, 2, 2), (2, 2, 4), (3, 1, 4), (0, 0, 8), ...

Here's how to cope with this problem: We're trying to count the number of ways to distribute 8 among the x_i. The 8 "units" to be distributed are indistinguishable balls, and each of the x_i is an urn which can "contain" balls. The urns are distinguishable because the x_i are distinguishable, e.g., the first two example solutions given above are different! So the answer is C(8+3–1, 8).

What if the x_i must be positive (not nonnegative)? The same, only we now require at least one ball per urn, that is, each x_i must be at least 1. The answer is C(7, 5).

Finally, what if it's $x_1 + x_2 + x_3 \leq 8$ (instead of = 8)? We handle this by introducing a "dummy" variable x_4 and counting solutions of $x_1 + x_2 + x_3 + x_4 = 8$; do you see why this gives the same answer?

Application: Runs

Suppose we have a string of 1s and 0s:

1 1 1 0 0 0 0 1 0 0 0 0 0 1 1

A *run* in such a sequence is a continuous string of the same number. The example sequence above has exactly five runs: First a run of three 1s, then a run of four 0s, then a run of a single 1, etc.

How many ways are there to arrange six 1s and nine 0s such that exactly five runs result?

There are two ways to make five runs: Either we start with a run of 1s as above (in which case we must also end with 1s) or we start (and end) with 0s.

Start with the first case: Let x_1, x_2, and x_3 be the number of 1s in the first, second, and third run. Each x_i must be positive and we must have $x_1+x_2+x_3=6$. There are C(5,3) possible solutions to this equation, that is, C(5,3) ways of making three runs from six 1s. Similarly, using $y_1+y_2=9$, there are C(8,7) ways of arranging nine 0s into two runs.

Multiplying, there are C(5,3)C(8,7) ways of making five runs starting with 1s. You do case 2, that is, count the number of ways of making five runs starting with 0s. What then should you do with these two numbers?