Math 114 Discrete Mathematics
Notes on combinatorics
D Joyce, Spring 2018

What is combinatorics? It's that part of discrete mathematics devoted to counting things. The things are often sets with some kind of structure, perhaps subsets of other sets with certain properites. There are various principles used in combinatorics and we'll look at some of the more basic ones including additive and multiplicative principles, permutations, combinations, binomial coefficients and Pascal's triangle.

The additive principle and the inclusion-exclusion principle. You've known the *additive principle* ever since you learned how to add. If you have two disjoint sets A and B, then the cardinality of their union is the sum of their cardinalities,

$$|A \cup B| = |A| + |B|.$$

In other words, if none of these things are those things, then the number of things altogether is the sum of the number of these things and the number of those things.

There's actually a proof for the additive principle, and that proof uses mathematical induction on the cardinality of B where the base case is when $|B| = 0$.

The additive principle generalizes to n sets. Given n pairwise disjoint sets $A_1, A_2, \ldots, A_n$, then

$$|A_1 \cup \cdots \cup A_n| = |A_1| + \cdots + |A_n|.$$

The inclusion-exclusion principle generalizes the additive principle to when the sets aren't disjoint.

In the case of two sets A and B, it says

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

With three sets, it says $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$.

For four sets, you can find the cardinality of their union by first adding the cardinalities of each of the four sets, then subtracting the cardinalities of all six of their double intersections, then adding the cardinalities of all four triple intersections, and finally subtracting the cardinality of the quadruple intersection. In summary, include each set, exclude each double intersection, include each triple intersection, and exclude the quadruple intersection.

Of course, the inclusion-exclusion principle extends to any finite number of sets.

The multiplicative principle, choices and stages, and tree diagrams. The basic multiplicative principle says that if you have m choices, and for each choice you have n second choices, then altogether you have mn choices.

One situation in which this occurs is when you take the Cartesian product of two sets A and B. The Cartesian product $A \times B$ consists of all ordered pairs (a, b) where $a \in A$ and $b \in B$. Then

$$|A \times B| = |A|\,|B|.$$

In many of our applications, however, what the second choices are depend on the first choice you make, so we're not looking at just Cartesian products of sets.

Suppose now that you're making choices in several stages and the number of choices m_i you can choose from at a stage i doesn't depend on previous choices you've made. Then the total number of outcomes for n stages $1, 2, \ldots, n$ is the product $m_1 m_2 \ldots m_n$.

A special case of this is the product of sets. Given finite sets $A_1, A_2, \ldots, A_n$, their product $A_1 \times A_2 \times \cdots \times A_n$ consists of ordered n-tuples $(a_1, a_2, \ldots, a_n)$ where each a_i belongs to the corresponding set A_i. In order to choose one of these ordered n-tuples, for the first stage you have a choice of choosing any one of the elements of A_1 to be a_1. The number of choices at stage 1 is the cardinality of A_1. For the second stage you have $|A_2|$ choices for a_2, and so forth. The multiplicative principle gives us the standard formula for the cardinality of the product

$$|A_1 \times A_2 \times \cdots \times A_n| = \prod_{i=1}^{n} |A_i|.$$

We'll use the multiplicative principle right away to count permutations and combinations.

Permutations. Suppose we want to count all the ways you can rearrange the letters in ROFL. There are a lot of them such as FROL, OLFR, etc. These rearrangements are called permutations. When choosing a permutation of ROFL, you have 4 choices for the first letter, 3 remaining choices for the second (since we cant choose the first letter again), 2 remaining choices for the third, and then the fourth is forced on us. Thus there are $4 \cdot 3 \cdot 2 \cdot 1 = 24$ choices altogether.

Counting permutations. Our main question, an easily answered one, is how may permutations are there on a set A of n elements? If n is small, say 4, then we can list all the permutations. Let's list all the rearrangements of *abcd*.

abcd	*bacd*	*cabd*	*dabc*
abdc	*badc*	*cadb*	*dacb*
acbd	*bcad*	*cbad*	*dbac*
acdb	*bcda*	*cbda*	*dbca*
adbc	*bdac*	*cdab*	*dcab*
adcb	*bdca*	*cdba*	*dcba*

There are 24 of them.

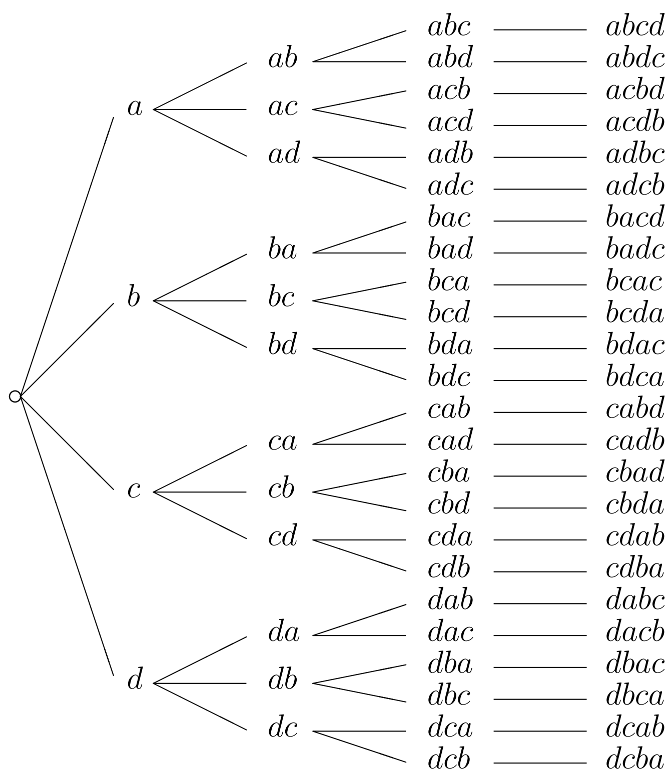
Even when n is not small, it's easy to determine how many permutations there are. We just use the multiplicative principle. In the first stage, choose one of the n elements to go first. In the second stage, there are $n-1$ remaining elements, and choose one of them to go second. At the next stage, chose one of the remaining $n-2$ elements to go next. And so forth until the last stage, when there's only one element left, so it goes last. Thus, the number of permutations of a set of n elements is

$$n(n-1)(n-2)\cdots 2 \cdot 1.$$

This last expression is usually abbreviated $n!$ and read "n factorial" or "factorial n" (except by some people who like to say "n shriek" or "n bang").

Thus, there are $4! = 24$ permutations of a set of 4 elements; $3! = 6$ permutations of a set of 3 elements; $2! = 2$ permutations of a set of 2 elements; $1! = 1$ permutations of a set of 1 element; and $0! = 1$ permutations of the empty set $\emptyset$. The last is because the unique function $\emptyset \to \emptyset$ is, by our definition, a permutation.

Tree diagrams. The stages in choosing a permutation can be illustrated in a *tree diagram*. When choosing a permutation of the four letters *abcd* there are four stages.



The first stage chooses one of the four letters to go first. That gives us our first branching of the tree at the left. After we've taken that branch, we'll be at one of the four *nodes* or *states* labelled a, b, c, or d. At this second stage, we choose a second letter that can't be the same as the first. In each case we have three choices this time, so we'll take one of the three branches to get to a state labelled by two letters. At the third stage, we've got two choices, so for each state there are two branches leading to a state labelled with three letters. At this state the last letter is determined, so there's only one branch to a *leaf* of the tree.

Sometimes variants of permutations come up. Here's one. Say you want to count the permutations of ROFLCOPTER. The two R's can't be distinguished, so exchanging them shouldn't count as a different permutation. If we said that there are 10! permutations, we would be doubly counting them because the R's aren't distinguishable. There are two O's, too, so the actual number of distinguishable permutations is 10! divided by 4.

Sterling's approximation for factorials. Sometimes you'll need to compute factorials of large numbers. Sterling's approximation helps. The factorial function $n!$ grows very fast with n. James Sterling (1692–1770) this approximation for factorials:

$$\boxed{n! \approx n^n e^{-n} \sqrt{2\pi n}}$$

This approximation is fairly good even for numbers as small as 10 where the approximation has an error of less than 1%. It's accuracy increases with n.

n	$n!$	approx	ratio
1	1	0.922137	1.084
2	2	1.91900	1.042
3	6	5.83621	1.028
4	24	23.5062	1.021
5	120	118.019	1.016
6	720	710.078	1.014
7	5040	4980.40	1.012
8	40320	39902.4	1.011
9	362880	359536	1.0093
10	3628800	3598690	1.0084
11	39916800	39615600	1.0076
12	479001600	475687000	1.0070

***k*-permutations.** Sometimes we don't want full permutations of a set of n elements, but just partial permutations. If $k \leq n$, a *k-permutation* is an ordered listing of just k elements of a set of n elements. For instance, the 3-permutations of *abcd* are these

abc	*bac*	*cab*	*dab*
abd	*bad*	*cad*	*dac*
acb	*bca*	*cba*	*dba*
acd	*bcd*	*cbd*	*dbc*
adb	*bda*	*cda*	*dca*
adc	*bdc*	*cdb*	*dcb*

while the 2-permutations are these

ab	*ba*	*ca*	*da*
ac	*bc*	*cb*	*db*
ad	*bd*	*cd*	*dc*

We can determine how many k-permutations of a set of n elements there are using the multiplicative principle. In the first stage, choose one of the n elements to go first. In the second stage, there are $n-1$ remaining elements, and choose one of them to go second. At the next stage, chose one of the remaining $n-2$ elements to go next. And so forth until the kth stage, when there are $n-k+1$ remaining elements. Thus, the number of k-permutations of a set of n elements is

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

There is no particular standard notation for the number of k-permutations of a set of n elements, but you'll see it denoted $(n)_k$ as in our text, nPk, P_k^n, and various other things. We'll probably just use $\dfrac{n!}{(n-k)!}$.

Definition of combinations and their relation to partial permutations. Combinations are related to partial permutations, but order is disregarded, as youll see.

A *combination* of size k from a set S of size n is just a subset of size k. It's more often it's called a k-subset when the size is specified.

A k-subset is related to k-permutations but they're not the same. A k-permutation is a listing of k distinct elements of S where the order of the elements in the listing is relevant. But for a k-subset, the elements are not listed in any particular order; that is, order doesn't matter.

Let's take an example. Let S be the 5-element set $S = \{a, b, c, d, e\}$. There are $5 \cdot 4 \cdot 3 = 60$ 3-permutations of S, but there are far fewer 3-subsets of S. For instance, one 3-subset is $\{a, b, c\}$. But this subset is associated to 6 of the 3-permutations, namely, *abc*, *acb*, *bac*, *bca*, *cab*, and *cba*. There are 6, of course, because there are $3! = 6$ full permutations of a set of 3 elements.

In general, each k-subset is associated to $k!$ of the k-permutations. Since there are $\dfrac{n!}{(n-k)!}$ of the k-permutations altogether, that implies that the number of k-subsets of a set of n elements is exactly $\dfrac{n!}{k!(n-k)!}$.

Binomial coefficients. That last expression is called a *binomial coefficient*, and it's denoted $\binom{n}{k}$, pronounced "n choose k".

$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

Thus, there are $\binom{n}{k}$ subsets of size k in a set of size n. Binomial coefficients are also called *combinations*, and an alternative notation for them is nCk.

Note that this definition is relevant so long as $0 \leq k \leq n$ and $n \geq 0$, where, as always, $0! = 1$.

The number $\binom{n}{k}$ of combinations of n things chosen k at a time is usually called a *binomial coefficient*. That's because they occur in the expansion of the n^{th} power of a binomial.

A *binomial* is a polynomial with two terms. Let's take the simplest binomial, $x + y$, and write up a table of its powers $(x+y)^n$ for the first few n.

$(x+y)^0 =$	1
$(x+y)^1 =$	$x+y$
$(x+y)^2 =$	$x^2 + 2xy + y^2$
$(x+y)^3 =$	$x^3 + 3x^2y + 3xy^2 + y^3$
$(x+y)^4 =$	$x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$
$(x+y)^5 =$	$x^5+5x^4y+10x^3y^2+10x^2y^3+5xy^4+y^5$

The coefficients in these polynomials, the powers of the binomial $x + y$, are the binomial coefficients. That's the *binomial theorem*.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k y^{n-k}.$$

To see why binomial coefficients count combinations, consider the coefficient 6 of x^2y^2. When you expand the product $(x+y)(x+y)(x+y)(x+y)$ you'll get a term x^2y^2 if you choose an x from exactly 2 of the 4 factors $x+y$, the y^2 coming from the remaining two factors. There are $\binom{4}{2} = 6$ ways of choosing 2 of the four factors, and each one contributes one x^2y^2, so the coefficient of x^2y^2 in the product will be $\binom{4}{2}$.

One important identity of the many important identities that hold for binomial coefficients is this one:

$$\binom{n}{k} = \binom{n}{n-k}$$

You can see why that's true in three different ways. First, they're both equal to $\dfrac{n!}{k!(n-k)!}$. Second as coefficients in the expansion of $(x+y)^n$, the coefficient of $x^k y^{n-k}$ is equal to the coefficient of $y^k x^{n-k}$. And third, each subset of k elements in a set of size n has a complement that has $n-k$ elements. The last reason is the best because it directly uses the meaning of $\binom{n}{k}$.

Pascal's triangle. We'll compute a few of these binomial coefficients. Then we'll show the recurrence relation

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1},$$

then display these binomial coefficients in a triangular table usually called *Pascal's triangle*.

Blaise Pascal (1623–1662) and Pierre de Fermat (1601–1665) studied these binomial coefficients in the context of probability in the 1600s. Their correspondence resulted in some of the first significant theory of probability and a systematic study of binomial coefficients. Because of Pascal's publication of their results, a particular arrangement of the binomial coefficients in a triangle is called *Pascal's triangle*. If you prefer, you can call it the *arithmetic triangle*. It was known in Europe for a couple of centuries before Pascal, and it was known much longer in Islamic mathematics, in India, and in China.

$$\begin{array}{ccccccccccc} &&&&&1&&&&& \\ &&&&1&&1&&&& \\ &&&1&&2&&1&&& \\ &&1&&3&&3&&1&& \\ &1&&4&&6&&4&&1& \\ 1&&5&&10&&10&&5&&1 \end{array}$$

The numbers along the sides are all 1s, and each entry in the middle is the sum of the two entries above it. These are just the binomial coefficients $\binom{n}{k}$ arranged in a table. By making it a triangle rather than a rectangle, you can see the two relationships

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \text{ and } \binom{n}{k} = \binom{n}{n-k}$$

more clearly. Here's the triangle again, but one less row of it so it fits on the page.

$$\begin{array}{ccccccccc} &&&&\binom{0}{0}&&&& \\ &&&\binom{1}{0}&&\binom{1}{1}&&& \\ &&\binom{2}{0}&&\binom{2}{1}&&\binom{2}{2}&& \\ &\binom{3}{0}&&\binom{3}{1}&&\binom{3}{2}&&\binom{3}{3}& \\ \binom{4}{0}&&\binom{4}{1}&&\binom{4}{2}&&\binom{4}{3}&&\binom{4}{4} \end{array}$$

There are lots of interrelations among these entries in Pascal's triangle, and we may have time to look at a couple of them. Note that in each row, n is fixed. Let's call that the n^{th} row; the top row is then the 0^{th} row. Note that the numbers in the n^{th} row sum to 2^n.

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

For instance, when $n = 5$, we have $1+5+10+10+5+1 = 32 = 2^5$. That's because these binomial coefficients tell us the number of subsets of various sizes of a set of n elements. Since there are 2^n subsets in all, they have to add up to 2^n.

There are different kinds of proofs you can give for these identities. You can prove this one using counting arguments or algebraic arguments.

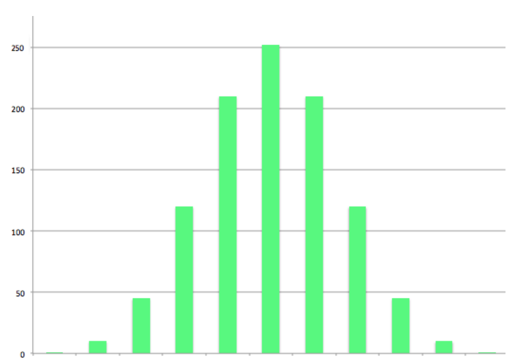
Counting proof: These binomial coefficients tell us the number of subsets of various sizes of a set of n elements. Since there are 2^n subsets in all, they have to add up to 2^n.

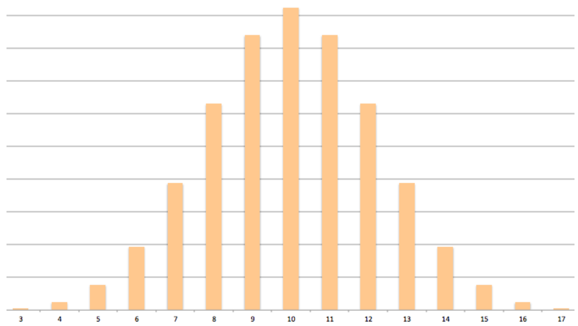
Algebraic proof: Use the binomial theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

with both x and y set to 1.

Graphic interpretation of binomial coefficients. It's interesting to look at a row of binomial coefficients displayed as bar chart, or histogram. Compare these graphs for $n = 10$ and $n = 20$.





They're on different scales, and only the center portion of the second graph is shown since the bars are too short to see outside the range shown. Their shapes are about the same, and, in the limit, give an important distribution of probability and statistics called the *normal distribution*, sometimes called the Gauss or Laplace-Gauss distribution, although it was first mentioned by De Moivre in 1733.

Math 114 Home Page at `http://math.clarku.edu/~djoyce/ma114/`