1 Overview

In this lecture, we begin studying the fundamentals of counting discrete objects. We will also give formalizations of rules and principles that we often use when counting objects.

2 Permutations and Combinations

In this section, we will introduce the foundations of combinatorics, the branch of mathematics that deals with counting. In particular, we will study permutations and combinations, their relevant formulas, and some basic identities involving these concepts.

Permuting n items: Suppose that we want to place n distinct items in a line, that is, we want to *order* or *permute* the items. The most fundamental principle used in counting is the following: the number of ways to permute n distinct items is the product of the first n positive integers, which is denoted by $n!$ ("n factorial"). This can be proven, somewhat informally, as follows: there are n choices for the first item in line, $(n-1)$ for the second, and so on, until there's only 1 choice left for the last item in line. Since the choices can be made in sequential order, the total number of possible permutations (orderings) is $n(n-1)\cdots 1 = n!$.

Permuting k of n items: Now instead of permuting all n items, suppose we only want to permute k of them (where k is a positive integer less than n). By the same argument above, the number of ways to do this is $n(n-1)\cdots(n-(k-1))$. Notice that there are k terms in this product, because each term corresponds to selecting an item to place next in the final ordering of k items. For convenience, notice that a more concise way of writing this product is the following:

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

Choosing k of n items: Now instead of permuting k items, suppose we only want to *choose* k items. In other words, we are finding a *subset* of size k, rather than a *sequence* of length k. We denote this value by $\binom{n}{k}$ ("n choose k"), and we often say a subset is a *combination*. In other words, a combination can be thought of as a permutation in which order doesn't matter.

As we just saw, there are $n!/(n-k)!$ permutations of length k. However, each subset of size k is represented as a permutation $k!$ times. Therefore, we can conclude that:

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

Alternatively, we can reason about these formulas as follows. Let k be any positive integer at most n. Then permuting n items is the same as choosing a subset of size k, ordering them, and then

ordering the remaining elements. This line of reasoning results in the following:

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

which is equivalent to the previous expression.

Remark: If $n = 0$, then the number of ways to order n items is somewhat ambiguous. Similarly, if $k = 0$, then the number of ways to choose k items is somewhat ambiguous. To deal with these ambiguities in a consistent way, we have the following conventions:

$$0! = 1 \qquad \text{and} \qquad \binom{n}{0} = 1.$$

2.1 Stars and Bars

The notion of combinations is fundamental to combinatorics. To better familiarize ourselves with combinations, we now look at one application known as "stars and bars".

The setup is the following: suppose there are three children c_1, c_2, c_3, and we must distribute 10 identical candies among these three children. Each child can receive any number of candies, including 0. For example, one possible distribution is $(4, 3, 3)$: in this case, c_1 receives 4 candies, c_2 receives 3, and c_3 receives 3. How many ways can we distribute the candies?

The key observation is the following: we can distribute the candies by arranging them in a line, and then placing two "bars" somewhere along the line. For example, the $(4, 3, 3)$ described above can be modeled by the following:

$$\bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \Big| \quad \bigstar \quad \bigstar \quad \bigstar \quad \Big| \quad \bigstar \quad \bigstar \quad \bigstar$$

Each $\bigstar$ represents a candy, and the two location of the two bars determines the distribution of the candies. Notice that the following distribution is also possible:
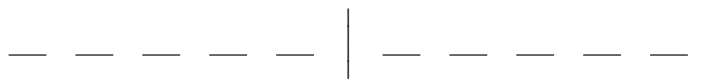
$$\Big| \quad \Big| \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar \quad \bigstar$$

The above diagram corresponds to the distribution $(0, 0, 10)$. In general, c_1 receives the candies left of the first bar, c_2 receives the candies between the two bars, and c_3 receives the candies right of the second bar.

So we can see that distributing candies is identical to choosing the location of the two bars. However, the number of ways to do this is *not* $\binom{11}{2}$, because this ignores the possibility of placing two bars next to each other. Instead, we should think of the process as follows: there are 12 empty slots, pictured below.

$$\text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—}$$
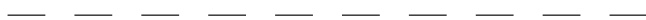
We must place bars in two of the slots, and the remaining 10 slots will then represent the 10 candies to distribute. For the distribution $(5, 5, 0)$, the diagram becomes the following:

$$\text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \Big| \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \text{—} \quad \Big|$$

From this perspective, it becomes clear that the number of distributions of 10 candies to 3 children is $\binom{12}{2}$. In general, if there are n candies and k children, then there are $n + k - 1$ slots, and we must place $k - 1$ bars. The remaining n candies, interspersed among the bars, represent a distribution. Thus, the number of distributions is $\binom{n+k-1}{k-1}$.

2.2 Combinatorial Identities

We now state a couple of basic identities involving combinations.

Fact 1. *Let n be a positive integer, and k be in $\{1, \ldots, n\}$. Then the following identities hold:*

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

The latter is known as Pascal's rule, *named after the mathematician Blaise Pascal.*

Remark: One proof of these identities is purely algebraic: if we simply use the formula

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

on every term, then a bit of algebraic manipulation proves the identity. However, in some sense, this proof is not very elegant because it does not take advantage of the combinatorial interpretation of the terms. Therefore, we now prove each identity by giving a combinatorial interpretation of both sides. Under these interpretations, the validity of the equalities becomes clear.

Proof. Let A be a set with n elements, where n is a positive integer. As discussed above, for each equality, we will give a combinatorial interpretation for both the left-hand side (LHS) and the right-hand side (RHS).

In the first equality, the LHS counts the number of subsets of A that contain exactly k items. *Selecting k items is equivalent to excluding $n - k$ items.* The number of ways to exclude $n - k$ items from S is precisely the RHS of the equality.

In the second equality, the LHS still counts the number of subsets of A that contain exactly k items. Since n is positive, we can fix a particular element of A which we denote by x. The set of subsets of size k can be partitioned into two sets: those that contain x, and those that do not. The number of subsets of size k that contain x is $\binom{n-1}{k-1}$. This is because once you choose x in the subset, you are left to make $k - 1$ choices out of the remaining $n - 1$ elements. On the other hand, the number of subsets of size k that exclude x is $\binom{n-1}{k}$. The logic behind this expression is that if you exclude x from the subset, then you are still left to make all the k choices, but only among the remaining $n - 1$ elements. Summing these two values gives the number of ways to choose *any* subset of size k, as desired. □

A pictorial representation of Pascal's rule (see Fact 1) is known as *Pascal's triangle*. The triangle is constructed as follows: the top row, which we consider row 0, contains a single 1. Each subsequent row starts and ends with 1, and the internal terms are obtained by summing the closest two terms in the previous row. The first 7 rows are pictured below, where n represents the row index:

$n=0$						1						
$n=1$					1		1					
$n=2$				1		2		1				
$n=3$			1		3		3		1			
$n=4$		1		4		6		4		1		
$n=5$	1		5		10		10		5		1	
$n=6$	1	6		15		20		15		6		1

Now consider the values in row n. We claim that these values are precisely, in order, the following:

$$\binom{n}{0} \quad \binom{n}{1} \quad \cdots \quad \binom{n}{n-1} \quad \binom{n}{n}.$$

The k-th term of this row (where $k \in \{0, 1, \ldots, n\}$) is precisely $\binom{n}{k-1}$—the proof of this is a simple induction argument that follows easily from Pascal's identity.

We now state and prove another well-known identity; this one gives a combinatorial interpretation of the coefficients of a binomial expression. Due to this connection, the values $\binom{n}{k}$ are often called the *Binomial coefficients*.

Theorem 2 (Binomial Theorem). *For all $n \in \mathbb{Z}^+$ and $a, b \in \mathbb{R}$,*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

In other words, for every $k \in \{0, 1, \ldots, n\}$, the coefficient of $a^{n-k}b^k$ in the expansion of $(a+b)^n$ is $\binom{n}{k}$.

As an example, let us consider the case when $n = 4$ and $k = 3$. Then the coefficient we seek is that of $a^1 b^3$. There are precisely four ways to obtain $a^1 b^3$ by expanding $(a+b)^4$: $abbb, babb, bbab$, and $bbba$. Each of these ways corresponds to choosing the location of 3 b's from four slots, and the number of ways to do this is precisely $\binom{4}{3}$.

Now we consider the general case. Let $k \in \{0, 1, \ldots n\}$, and consider the term $a^{n-k}b^k$. This term is the product of $(n-k)$ a's and k b's. Furthermore, notice that $(a+b)^n$ is simply

$$(a+b)(a+b) \cdots (a+b),$$

which is the product of n copies of $(a+b)$. Upon expansion of this product, we obtain the sum of a $n+1$ terms, and each term is the product of n variables, distributed as a's and b's.

Thus, we can think of "building" the coefficient of $a^{n-k}b^k$ as selecting k of the $(a+b)$ terms in $(a+b)^n$ that will contribute a b. The number of ways to do this is precisely $\binom{n}{k}$, as desired.

Corollary 3. *For all $n \in \mathbb{Z}^+$, the following equality holds:*

$$2^n = \sum_{k=0}^{n} \binom{n}{k}$$

In other words, the sum of the values in the n-th row of Pascal's triangle is 2^n.

Remark: The above identity simply follows by $a = b = 1$ in Theorem 2. We give an alternate proof below that relates this identity to the set of subsets of a set.

Proof. Let A be a set with n elements, and let A_k denote the subset of the power set 2^A containing the subsets of A of size k. Then the sets $A_0, A_1, \ldots, A_n$ partition 2^A, which means the following equalities hold:

$$2^n = |2^A| = |A_0| + |A_1| + \cdots + |A_n|$$
$$= \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = \sum_{k=0}^{n} \binom{n}{k}.$$

In other words, every subset of A has a size in $\{0, 1, \ldots, n\}$, so to count the number of subsets of A, we can count the number of subsets of each size over all possible sizes. $\square$

3   Cardinality Rules and Principles

In this section, we will see the formalization of counting strategies that we often take for granted: the product rule, the sum rule, and the pigeonhole principle.

3.1   The Product Rule

Before stating this rule, we first introduce some notation. For any set S and positive integer k, we let S^k denote the set of k-length sequences whose elements are in S, i.e.,

$$S^k = \{(s_1, s_2, \ldots, s_k) : \forall i \in \{1, 2, \ldots, k\}. s_i \in S\}.$$

Notice that $S^2 = S \times S$, but $S^3 \neq (S \times S) \times S$ because the elements of S^3 are of the form (a, b, c), while the elements of $(S \times S) \times S$ have the form $((a, b), c)$. However, for notational convenience, we often remove the internal parentheses and write (a, b, c) instead. Using this convention, the set S^k is the Cartesian product of S with itself k times.

The product rule states the following: *If A and B are finite sets, then $|A \times B| = |A| \cdot |B|$.* This rule allows us to count the number of different ways to "combine" choices. For example, suppose we must create a password that contains five letters of the alphabet (all lowercase), followed by three digits (i.e., 0-9). Let $A = \{a, b, c, \ldots\}$ denote the set of letters, and $B = \{0, 1, \ldots, 9\}$ denote the set of digits. Then to create a valid password, we must select five elements of A (order matters), as well as three elements of B (order matters as well).

Thus, the set P of valid passwords is essentially the set $A^5 \times B^3$: for example, "*apple*123" is a valid password, and $(a, p, p, l, e, 1, 2, 3)$ is an element of P (with internal parentheses removed). By the product rule, we have $|P| = |A^5 \times B^3| = |A^5| \cdot |B^3| = 26^5 \cdot 10^3$. Thus, there are $26^5 \cdot 10^3$ passwords that satisfy the given conditions.

3.2   The Principle of Inclusion-Exclusion

The principle of inclusion-exclusion (PIE) states the following: *If A and B are finite sets, then $|A \cup B| = |A| + |B| - |A \cap B|$.* Note that if $A \cap B = \emptyset$, then PIE states $|A \cup B| = |A| + |B|$; this special case is known as the *sum rule*. The name of this principle comes from the fact that we must "exclude" the elements of $A \cap B$ after "including" each of them twice—once as an element of A, and once as an element of B.

Inclusion-exclusion allows us to count the number of objects that satisfy some property or another, without over-counting the objects that satisfy both. For example, suppose 15 students in the class own a cat, 20 own a dog, and 10 own both. Then the number of students that own a cat or a dog (or both) is not $15 + 20 = 35$, but rather, $15 + 20 - 10 = 25$ because the 10 people that own both are included in both the 15 cat owners and the 20 dog owners.

3.3 The Pigeonhole Principle

A pigeonhole is a small box in which a single pigeon can stand. If there are n pigeonholes and over n pigeons, then clearly, some pigeonhole must accommodate at least 2 pigeons—this idea is known as the *pigeonhole principle*. More formally, the pigeonhole principle states the following: *If A and B are finite sets such that $|A| > |B|$ and $f : A \to B$ is a total function, then there exist $a_1, a_2 \in A$ such that $a_1 \neq a_2$ and $f(a_1) = f(a_2)$.* Thus, we can think of A as the set of pigeons, B as the set of pigeonholes, and f as the function that matches each pigeon with a pigeonhole. If $|A| > |B|$ then some pigeonhole must accommodate at least 2 pigeons, which is precisely what the principle states. The pigeonhole principle sounds quite obvious, yet when applied correctly, it can be surprisingly powerful.

Example 1: Consider a subset S of $\{1, 2, \ldots, 10\}$. We claim that if $|S| \geq 6$, then there exist two distinct elements of S such that one is a factor of the other. This proposition can be verified by checking all $\binom{10}{6}$ subsets, but we can provide a more elegant proof using the pigeonhole principle.

Notice that every positive integer n can be written as $2^k \cdot q(n)$, where k is a nonnegative integer and $q(n)$ is odd. Now consider writing the elements of S in this way. There are only 5 possible values of $q(n)$ (i.e., $\{1, 3, 5, 7, 9\}$), so there exists some odd integer q and $a, b \in S$ such that $a = 2^\alpha q$ and $b = 2^\beta q$, where α, β are distinct nonnegative integers. If $\alpha < \beta$ then a is a factor of b, and if $\alpha > \beta$, then b is a factor of a. In this example, we can think of the "odd component" of the elements of S as the pigeons, and the set $\{1, 3, 5, 7, 9\}$ as the pigeonholes.

4 Summary

In this lecture, we studied the fundamentals of combinatorics. We proved several combinatorial identities and saw applications of cardinality rules and the pigeonhole principle.