# Math 235 Fall 2024, Lecture 8 stenogram: Enumerative combinatorics

**website:** `https://www.cip.ifi.lmu.de/~grinberg/t/24f`

# 1. An enumerative combinatorics toolbox

Today I will talk about **enumerative combinatorics**, i.e., computing cardinalities of finite sets, also known as **counting**.

This is Chapter 7 of the notes, and there is more in the notes, including more details, more examples, and more references.

## 1.1. Basic rules for counting

- **Bijection principle:** Two sets $X$ and $Y$ satisfy $|X| = |Y|$ if and only if there exists a bijection $f : X \to Y$.
- **Sum rule:** For $k$ finite sets $S_1, S_2, \ldots, S_k$, we always have
$$|S_1 \cup S_2 \cup \cdots \cup S_k| \leq |S_1| + |S_2| + \cdots + |S_k|.$$
Equality holds if and only if the sets $S_1, S_2, \ldots, S_k$ are disjoint.
- **Product rule:** For $n$ finite sets $A_1, A_2, \ldots, A_n$, we always have
$$|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdot \cdots \cdot |A_n|.$$
(For $n = 0$, this is saying $|\{()\}| = 1$.)
In particular, for any finite set $A$, we have
$$|A^n| = |A|^n.$$
- **Difference rule:** If $B$ is a subset of a finite set $A$, then
$$|A \setminus B| = |A| - |B|.$$

## 1.2. Notations

- If $n \in \mathbb{N}$, then $[n]$ shall mean the $n$-element set $\{1, 2, \ldots, n\}$. In particular, $[1] = \{1\}$ and $[0] = \varnothing$.
- If $\mathcal{A}$ is a logical statement, then $[\mathcal{A}]$ means its **truth value**, defined by
$$[\mathcal{A}] = \begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false.} \end{cases}$$
For example, $[2+2=4] = 1$ and $[2+2=5] = 0$.
- The symbol "#" stands for "number".

## 1.3. Fundamental counting results

Not all counting problems have explicit answers, but many of the most elementary ones do. Here are some:

**Theorem 1.3.1.** Let $n \in \mathbb{N}$. Let $S$ be an $n$-element set. Let $k \in \mathbb{R}$. Then,

$$(\# \text{ of } k\text{-element subsets of } S) = \binom{n}{k}.$$

*Proof.* Induct on $n$ for a straightforward proof. See Theorem 4.3.12 in the notes for referenes. □

**Theorem 1.3.2.** Let $n \in \mathbb{N}$. Let $S$ be an $n$-element set. Then,

$$(\# \text{ of subsets of } S) = 2^n.$$

*Proof.* We denote the $n$ elements of $S$ by $s_1, s_2, \ldots, s_n$ (in some order). Now there is a bijection

$$\begin{aligned} \{\text{subsets of } S\} &\to \{0,1\}^n, \\ I &\mapsto ([s_1 \in I], [s_2 \in I], \ldots, [s_n \in I]). \end{aligned}$$

The inverse map sends each bitstring $(a_1, a_2, \ldots, a_n) \in \{0,1\}^n$ to the subset $\{i \in [n] \mid a_i = 1\}$. Details in the notes. □

**Theorem 1.3.3.** Let $n \in \mathbb{N}$. A **composition** of $n$ shall mean a tuple of positive integers whose sum is $n$. Then,

$$(\# \text{ of compositions of } n) = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 1, & \text{if } n = 0. \end{cases}$$

**Example 1.3.4.** The compositions of 3 are

$$(1,1,1), \qquad (1,2), \qquad (2,1), \qquad (3).$$

There are 4 of them, as the theorem predicts ($2^{3-1} = 4$).

*Proof of the theorem.* The case $n = 0$ is obvious (the only composition of 0 is the empty tuple ()). So we WLOG assume that $n > 0$. Hence, we must prove that

$$(\# \text{ of compositions of } n) = 2^{n-1}.$$

The preceding theorem says that

$$(\# \text{ of subsets of } [n-1]) = 2^{n-1}.$$

So we must prove that these two #s are equal. We can achieve this by constructing a bijection

$$f : \{\text{compositions of } n\} \to \{\text{subsets of } [n-1]\}.$$

We define this bijection $f$ as follows:

$$\begin{aligned} f(a_1, a_2, \ldots, a_k) &= \{a_1,\ a_1 + a_2,\ a_1 + a_2 + a_3,\ \ldots,\ a_1 + a_2 + \cdots + a_{k-1}\} \\ &= \{a_1 + a_2 + \cdots + a_i \mid i \in [k-1]\}. \end{aligned}$$

The inverse map $f^{-1}$ sends any given subset $I = \{i_1 < i_2 < \cdots < i_{k-1}\} \subseteq [n-1]$ to the composition

$$(i_1 - i_0,\ i_2 - i_1,\ i_3 - i_2,\ \ldots,\ i_{k-1} - i_{k-2},\ i_k - i_{k-1}),$$

where we set $i_0 := 0$ and $i_k := n$. Details in the notes. □

**Theorem 1.3.5.** Let $A$ and $B$ be two finite sets. Set $m = |A|$ and $n = |B|$. Then,

$$(\text{\# of maps from } A \text{ to } B) = n^m.$$

*Proof.* Idea: A map $f : A \to B$ can be determined by specifying all its values. That is, for each $a \in A$, we choose the value $f(a) \in B$. There are $n$ options for each of these values, and these choices are independent. So the total # of possibilities is $\underbrace{nn \cdots n}_{m \text{ times}} = n^m$.

More formally: Let the elements of $A$ be $a_1, a_2, \ldots, a_m$. Then, there is a bijection

$$\begin{aligned} \{\text{maps from } A \text{ to } B\} &\to B^m, \\ f &\mapsto (f(a_1), f(a_2), \ldots, f(a_m)). \end{aligned}$$

Thus, by the bijection principle,

$$|\{\text{maps from } A \text{ to } B\}| = |B^m| = |B|^m = n^m.$$

□

**Theorem 1.3.6.** Let $A$ and $B$ be two finite sets. Set $m = |A|$ and $n = |B|$. Then,

$$\begin{aligned} (\text{\# of injective maps from } A \text{ to } B) &= n(n-1)(n-2)\cdots(n-m+1) \\ &= m! \cdot \binom{n}{m}. \end{aligned}$$

*Proof.* Idea: A map $f : A \to B$ can be determined by specifying all its values. That is, for each $a \in A$, we choose the value $f(a) \in B$. However, to make $f$ injective, we must choose distinct values, i.e., we must ensure that no value is chosen more than once. So there are $n$ options for the first value, $n-1$ options for the second, $n-2$ options for the third, and so on, for a total of $m$ choices. So the total # of possibilities is

$$n(n-1)(n-2)\cdots(n-m+1) = m! \cdot \binom{n}{m}.$$

Formally... this is more complicated. The easiest way to formalize this can be found in [19fco, §2.4.2]; this proceeds by induction on $m$. Alternatively, the above "choices" argument can be crystallized into a general principle, formalized and proved. Let me state the principle, but not prove it here: □

**Theorem 1.3.7** (dependent product rule)**.** Consider a situation in which you must make $n$ decisions (in order). Assume that

- you have $a_1$ options in decision 1;
- you have $a_2$ options in decision 2 (no matter what option you chose in decision 1);
- you have $a_3$ options in decision 3;
- and so on.

Then, the total # of possibilities to make all these choices is $a_1 a_2 \cdots a_n$.

## 1.4. Some exercises

**Exercise 1** (Putnam 1985/A1)**.** Let $n \in \mathbb{N}$. Find the # of all triples $(A, B, C)$ of subsets of $[n]$ such that $A \cap B \cap C = \varnothing$.

*Solution.* We construct such a triple $(A, B, C)$ as follows: For each $i \in [n]$, we decide which of the three subsets it goes into. The only available options are "none", "$A$", "$B$", "$C$", "$A$ and $B$", "$A$ and $C$" and "$B$ and $C$" (but not "all three", since we want $A \cap B \cap C = \varnothing$). This makes 7 options per element, and thus $7^n$ possibilities altogether. So the # is $7^n$.

For the next exercise, we introduce some terminology:

**Definition 1.4.1.** A tuple $(a_1, a_2, \ldots, a_k)$ is said to be

- **injective** if all $a_1, a_2, \ldots, a_k$ are distinct;

- **Smirnov** (or **Carlitz** or **non-stuttering**) if it satisfies $a_i \neq a_{i+1}$ for all $i \in [k-1]$;
- **cyc-Smirnov** if it is Smirnov and also satisfies $a_k \neq a_1$.

**Exercise 2.** Let $n, k \in \mathbb{N}$. Let $A$ be an $n$-element set.
**(a)** How many injective $k$-tuples are there?
**(b)** How many Smirnov $k$-tuples are there (for $k > 0$)?
**(c)** How many cyc-Smirnov $k$-tuples are there (for $k > 0$)?

*Solution.* **(a)** The # of injective $k$-tuples is $\dbinom{n}{k} \cdot k!$.
Indeed, there is a bijection

$$\begin{aligned} \{\text{injective maps from } [k] \text{ to } A\} &\to \{\text{injective } k\text{-tuples}\}, \\ f &\mapsto (f(1), f(2), \ldots, f(k)). \end{aligned}$$

Thus, by the bijection principle, we reduce our problem to the problem of counting injective maps from $[k]$ to $A$, but for this we know the answer.

**(b)** The # of Smirnov $k$-tuples is $n(n-1)^{k-1}$.

Indeed, we can construct a Smirnov $k$-tuple $(a_1, a_2, \ldots, a_k)$ by first choosing $a_1$ (there are $n$ choices), then choosing $a_2$ (there are $n-1$ choices to ensure $a_1 \neq a_2$), then choosing $a_3$ (there are $n-1$ choices to ensure $a_2 \neq a_3$), and so on.

**(c)** The # of cyc-Smirnov $k$-tuples is ???

Let us denote this # by $c(n,k)$. We try to get a recursion for it.

We have

$$\begin{aligned} c(n,k) &= (\text{\# of cyc-Smirnov } k\text{-tuples}) \\ &= \underbrace{(\text{\# of Smirnov } k\text{-tuples})}_{=n(n-1)^{k-1}} \\ &\qquad - \underbrace{(\text{\# of Smirnov-but-non-cyc-Smirnov } k\text{-tuples})}_{\substack{=(\text{\# of Smirnov } k\text{-tuples } (a_1,a_2,\ldots,a_k) \text{ with } a_k=a_1) \\ =(\text{\# of cyc-Smirnov } (k-1)\text{-tuples}) \\ \text{(because there is a bijection} \\ \text{from } \{\text{Smirnov } k\text{-tuples } (a_1,a_2,\ldots,a_k) \text{ with } a_k=a_1\} \\ \text{to } \{\text{cyc-Smirnov } (k-1)\text{-tuples}\}, \\ \text{which just removes the last entry)}}} \\ &= n(n-1)^{k-1} - \underbrace{(\text{\# of cyc-Smirnov } (k-1)\text{-tuples})}_{=c(n,k-1)} \\ &= n(n-1)^{k-1} - c(n,k-1). \end{aligned}$$

This is a recursion that we can solve by plugging it into itself:

$$\begin{aligned}
c(n,k) &= n(n-1)^{k-1} - \underbrace{c(n,k-1)}_{=n(n-1)^{k-2}-c(n,k-2)} \\
&= n(n-1)^{k-1} - n(n-1)^{k-2} + \underbrace{c(n,k-2)}_{=n(n-1)^{k-3}-c(n,k-3)} \\
&= \cdots \\
&= n(n-1)^{k-1} - n(n-1)^{k-2} + n(n-1)^{k-3} - n(n-1)^{k-4} \pm \cdots \\
&= (n-1)^k + (-1)^k (n-1).
\end{aligned}$$

So the answer is $(n-1)^k + (-1)^k (n-1)$. It is easy to prove this by induction using our recursion $c(n,k) = n(n-1)^{k-1} - c(n,k-1)$.

## 1.5. Permutations

Recall that a **permutation** of a set $X$ means a bijection from $X$ to $X$. The following are easy:

**Theorem 1.5.1.** Let $n \in \mathbb{N}$. For any two $n$-element sets $U$ and $V$, we have

$$(\text{\# of bijections from } U \text{ to } V) = n!.$$

In particular, for any $n$-element set $X$, we have

$$(\text{\# of permutations of } X) = n!.$$

**Definition 1.5.2.** Let $n \in \mathbb{N}$. The set of all permutations of $[n]$ is called the **$n$-th symmetric group**, and is denoted by $S_n$.

This is really a group in the sense of abstract algebra, i.e., it is closed under composition ($f \circ g \in S_n$ whenever $f, g \in S_n$) and under taking inverses.

**Exercise 3.** Let $n \geq 2$. How many permutations $\sigma \in S_n$ satisfy $\sigma(1) > \sigma(2)$ ?

*First solution.* We construct such a permutation $\sigma$ as follows:

- First we choose the two values $\sigma(1)$ and $\sigma(2)$. The # of options for this is $\binom{n}{2}$, since we are really choosing the 2-element subset $\{\sigma(1), \sigma(2)\}$ of $[n]$. The condition $\sigma(1) > \sigma(2)$ dictates that $\sigma(1)$ must be the largest element of this subset, and $\sigma(2)$ its smallest element.

- Then we choose the remaining values $\sigma(3), \sigma(4), \ldots, \sigma(n)$. The # of options for this is $(n-2)(n-3)\cdots 1 = (n-2)!$.

So the total # of permutations $\sigma$ is $\dbinom{n}{2} \cdot (n-2)!$.

*Second solution.* A symmetry argument: Among the permutations $\sigma \in S_n$, equally many satisfy $\sigma(1) > \sigma(2)$ as satisfy $\sigma(1) < \sigma(2)$ (since we can get from one to the other by swapping the values $\sigma(1)$ and $\sigma(2)$). So the # of permutations $\sigma \in S_n$ that satisfy $\sigma(1) > \sigma(2)$ must be $\dfrac{n!}{2}$ (since the total # of permutations $\sigma \in S_n$ is $n!$).

As a nice bonus, we have obtained the identity

$$\underbrace{\binom{n}{2} \cdot (n-2)!}_{\text{our first answer}} = \underbrace{\frac{n!}{2}}_{\text{our second answer}} \qquad \text{for each } n \geq 2.$$

**Definition 1.5.3.** Let $X$ be a set. Let $f : X \to X$ be a map.

**(a)** A **fixed point** of $f$ means an $x \in X$ such that $f(x) = x$.

**(b)** We let $\operatorname{Fix} f$ denote the set of all fixed points of $f$.

**Exercise 4** (IMO 1987 problem 1)**.** Let $n$ be a positive integer. Prove that

$$\sum_{w \in S_n} |\operatorname{Fix} w| = n!.$$

In other words, prove that on average, a random permutation of $[n]$ has 1 fixed point.

*Solution.* We use the language of probabilities. We let $\operatorname{E}(X)$ denote the expected value of a random variable $X$, whereas $\Pr(A)$ denotes the probability of an event $A$. Note that

$$\Pr(A) = \operatorname{E}([A]) \qquad \text{for an event } A.$$

We are looking for the average # of fixed points of a random permutation of $[n]$. In other words, we are looking for $\operatorname{E}(|\operatorname{Fix} w|)$ where $w \in S_n$ is uniformly random.

Now a basic property of expected values comes handy: If $X_1, X_2, \ldots, X_k$ are any $k$ random variables, then

$$\operatorname{E}(X_1 + X_2 + \cdots + X_k) = \operatorname{E}(X_1) + \operatorname{E}(X_2) + \cdots + \operatorname{E}(X_k).$$

This is known as **linearity of expectation**. We can apply it by decomposing $|\operatorname{Fix} w|$ as

$$|\operatorname{Fix} w| = [1 \in \operatorname{Fix} w] + [2 \in \operatorname{Fix} w] + \cdots + [n \in \operatorname{Fix} w].$$

We get

$$\begin{aligned}
\operatorname{E}(|\operatorname{Fix} w|) &= \operatorname{E}([1 \in \operatorname{Fix} w] + [2 \in \operatorname{Fix} w] + \cdots + [n \in \operatorname{Fix} w]) \\
&= \operatorname{E}([1 \in \operatorname{Fix} w]) + \operatorname{E}([2 \in \operatorname{Fix} w]) + \cdots + \operatorname{E}([n \in \operatorname{Fix} w]) \\
&\qquad \text{(by linearity of expectation)} \\
&= \Pr(1 \in \operatorname{Fix} w) + \Pr(2 \in \operatorname{Fix} w) + \cdots + \Pr(n \in \operatorname{Fix} w) \\
&\qquad \left(\text{since } \operatorname{E}([A]) = \Pr(A) \text{ for any event } A\right).
\end{aligned}$$

Now, how can we compute the probabilities on the RHS?

Fix $i \in [n]$, and try to compute $\Pr(i \in \operatorname{Fix} w)$.

*First way:* The # of permutations $w \in S_n$ that fix $i$ is $(n-1)!$, since choosing such a permutation $w$ means choosing its remaining $n-1$ values. That is, $\Pr(i \in \operatorname{Fix} w) = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$.

*Second way:* We note that $\Pr(i \in \operatorname{Fix} w) = \Pr(w(i) = i)$. But $w(i)$ must always be one of the $n$ numbers $1, 2, \ldots, n$, and moreover, each of these $n$ numbers is equally likely (since there are bijections between these possibilities). So all the probabilities

$$\Pr(w(i) = 1), \quad \Pr(w(i) = 2), \quad \ldots, \quad \Pr(w(i) = n)$$

are equal, but they add up to 1. So each of them equals $\dfrac{1}{n}$. In particular, $\Pr(w(i) = i) = \dfrac{1}{n}$, so that $\Pr(i \in \operatorname{Fix} w) = \dfrac{1}{n}$.

Either way, we have shown that $\Pr(i \in \operatorname{Fix} w) = \dfrac{1}{n}$. Now,

$$\begin{aligned}
\operatorname{E}(|\operatorname{Fix} w|) &= \Pr(1 \in \operatorname{Fix} w) + \Pr(2 \in \operatorname{Fix} w) + \cdots + \Pr(n \in \operatorname{Fix} w) \\
&= \underbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}_{n \text{ times}} = 1.
\end{aligned}$$

In other words, on average, a permutation $w \in S_n$ has 1 fixed point.

See the notes for details.

## 1.6. Double counting

Often, a counting problem can be solved in two ways, yielding two results. As a consequence, the two results are equal, even if this is not immediately obvious from their looks. For instance, we obtained

$$\binom{n}{2} \cdot (n-2)! = \frac{n!}{2}$$

by solving a counting problem and obtaining both sides as answers.

This identity is easy to check by hand, but often you can find nontrivial identities this way. Sometimes, when proving an identity (usually a binomial identity), you can interpret its LHS and its RHS as two answers to one and the same counting problem, and thus prove the identity. This technique is called **double counting**.

For example, how would you prove

$$2^n = \sum_{k=0}^{n} \binom{n}{k} \qquad \text{for all } n \in \mathbb{N}$$

by double counting? You count subsets of $[n]$. The LHS counts them directly, while the RHS counts them by size. Because the two sides answer the same problem, they must be equal, so the identity is (re)proved.

Here are two more interesting examples for this technique:

**Proposition 1.6.1** (Chu–Vandermonde identity for nonnegative integers)**.** Let $n, x, y \in \mathbb{N}$. Then,

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

**Proposition 1.6.2** (Trinomial revision formula for nonnegative integers)**.** Let $n, a, b \in \mathbb{N}$. Then,

$$\binom{n}{a}\binom{a}{b} = \binom{n}{b}\binom{n-b}{a-b}.$$