Math 222 Fall 2022, Lecture 23: Binomial coefficients

website: `https://www.cip.ifi.lmu.de/~grinberg/t/22fco`

2. Binomial coefficients (cont'd)

2.10. Multinomial coefficients (cont'd)

2.10.1. Definition and formulas (cont'd)

What gave multinomial coefficients their name is the following theorem (which generalizes the binomial formula):

> **Theorem 2.10.5** (multinomial formula)**.** Let $x_1, x_2, \ldots, x_k$ be any numbers. Let $n \in \mathbb{N}$. Then,
>
> $$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k; \\ a_1 + a_2 + \cdots + a_k = n}} \binom{n}{a_1, a_2, \ldots, a_k} x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}.$$

Proof. Either a straightforward induction on n, or a straightforward induction on k, or expand and use the combinatorial interpretation of multinomial coefficients (Proposition 2.10.8 below). References can be found in the 2019 notes (proof of Theorem 2.12.16). □

2.10.2. Counting maps that take each value a given number of times

Here is a combinatorial interpretation of multinomial coefficients:

> **Proposition 2.10.6.** Let $n \in \mathbb{N}$ and $n_1, n_2, \ldots, n_k \in \mathbb{N}$ be such that $n_1 + n_2 + \cdots + n_k = n$. Then, the # of maps $f : [n] \to [k]$ that satisfy
>
> $$(\text{\# of } a \in [n] \text{ such that } f(a) = i) = n_i \qquad \text{for each } i \in [k]$$
>
> is $\binom{n}{n_1, n_2, \ldots, n_k}$.

> **Example 2.10.7.** Applying this proposition to $n = 7$ and $k = 3$ and $(n_1, n_2, n_3) = (2, 3, 2)$ yields that the # of maps $f : [7] \to [3]$ that take the value 1 two times, the value 2 three times, and the value 3 two times is
>
> $$\binom{7}{2, 3, 2} = \frac{7!}{2! \cdot 3! \cdot 2!} = 210.$$

An example of such a map is (written in two-line notation)

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 1 & 2 & 2 & 2 & 3 & 3 \end{pmatrix}.$$

Another is

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 1 & 3 & 1 & 2 & 2 \end{pmatrix}.$$

We are not going to list the remaining 208 such maps.

Proof of Proposition 2.10.6. We need to count the maps $f : [n] \to [k]$ that satisfy

$$(\# \text{ of } a \in [n] \text{ such that } f(a) = i) = n_i \qquad \text{for each } i \in [k].$$

We can construct such a map in the following way:

1. We choose the set $\{a \in [n] \mid f(a) = 1\}$. This must be an n_1-element subset of $[n]$, so we have $\dbinom{n}{n_1}$ many options for it.

2. We choose the set $\{a \in [n] \mid f(a) = 2\}$. This must be an n_2-element subset of
$$[n] \setminus \{a \in [n] \mid f(a) = 1\},$$
so we have $\dbinom{n - n_1}{n_2}$ many options for it[1].

3. We choose the set $\{a \in [n] \mid f(a) = 3\}$. This must be an n_3-element subset of
$$[n] \setminus (\{a \in [n] \mid f(a) = 1\} \cup \{a \in [n] \mid f(a) = 2\}),$$
so we have $\dbinom{n - n_1 - n_2}{n_3}$ many options for it[2].

4. And so on.

By the dependent product rule, we thus see that the total # of such maps is

$$\begin{aligned} & \binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2-\cdots-n_{k-1}}{n_k} \\ &= \binom{n}{n_1, n_2, \ldots, n_k} \qquad \left(\begin{array}{c}\text{by Proposition 2.10.4 } \textbf{(a)} \\ \text{from Lecture 22}\end{array}\right). \end{aligned}$$

So we are done.

(See Proposition 2.12.5 in the 2019 notes for more details.) □

Let me restate Proposition 2.10.6 using tuples instead of maps:

[1] since $[n] \setminus \{a \in [n] \mid f(a) = 1\}$ is an $(n - n_1)$-element set

[2] since $[n] \setminus (\{a \in [n] \mid f(a) = 1\} \cup \{a \in [n] \mid f(a) = 2\})$ is an $(n - n_1 - n_2)$-element set

Proposition 2.10.8. Let $n \in \mathbb{N}$ and $n_1, n_2, \ldots, n_k \in \mathbb{N}$ be such that $n_1 + n_2 + \cdots + n_k = n$. Then, the # of n-tuples $(u_1, u_2, \ldots, u_n) \in [k]^n$ that satisfy

$$(\text{\# of } a \in [n] \text{ such that } u_a = i) = n_i \qquad \text{for each } i \in [k]$$

is $\dbinom{n}{n_1, n_2, \ldots, n_k}$.

Proof. This follows from Proposition 2.10.6, using the standard "encode maps as tuples" bijection. (Again, details can be found in the 2019 notes – Proposition 2.12.7 to be specific.) □

2.10.3. Counting anagrams

An equivalent version of Proposition 2.10.6 can be formulated using the notion of an "anagram". Let us define this first:

Definition 2.10.9. Let $n \in \mathbb{N}$. Let α be an n-tuple (of any objects).

An **anagram** of α shall mean an n-tuple that can be obtained from α by permuting the entries.

In other words, if $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, then an **anagram** of α means an n-tuple of the form

$$\left(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(n)}\right), \qquad \text{where } \sigma \text{ is a permutation of } [n].$$

For example, the 6-tuple $(1, 3, 2, 4, 4, 2)$ is an anagram of $(2, 2, 1, 4, 3, 4)$, because the former tuple can be obtained from the latter by permuting the entries (namely, permuting them in such a way that the entry 1 moves to the front, the two entries 2 move to the third and sixth positions, the entry 3 moves to the second position, and the two entries 4 move to the fourth and fifth positions).

Anagrams are often called "permutations" in mathematics. I prefer calling them "anagrams" in order to reserve the word "permutation" for a bijective map from a set to itself. So a permutation for me means (roughly speaking) a **way** to permute elements, not a **result** of permuting elements. Distinct permutations of $[n]$ may lead to the same anagram of α (if α has equal entries), so it is important to keep the two apart, and I do this by giving them different names.

Example 2.10.10. How many anagrams does the word "anagram" have?

We treat a word as a tuple of letters, so the word "anagram" is the 7-tuple (a, n, a, g, r, a, m).

We note that an anagram of this 7-tuple is just a 7-tuple that contains the letter a three times, the letter n once, the letter g once, the letter r once, and the letter m once. The # of such tuples can be obtained from Proposition

2.10.6 (after renaming the letters a, n, g, r, m as $1, 2, 3, 4, 5$, respectively). We conclude that the # of such tuples is $\binom{7}{3,1,1,1,1} = \frac{7!}{3! \cdot 1! \cdot 1! \cdot 1! \cdot 1!} = 840$.

By the same logic, we obtain the following general theorem:

Theorem 2.10.11. Let $n \in \mathbb{N}$ and $n_1, n_2, \ldots, n_k \in \mathbb{N}$ be such that $n_1 + n_2 + \cdots + n_k = n$.

Let α be an n-tuple that contains n_1 many 1's, n_2 many 2's, ..., n_k many k's.

Then, the # of distinct anagrams of α is $\binom{n}{n_1, n_2, \ldots, n_k}$.

Proof. See Proposition 2.12.13 in the 2019 notes (but this is just the argument from Example 2.10.10, generalized and formalized). □

2.11. Odds and ends

Here is a curious counting puzzle that can be solved with what we have seen so far, and is actually pretty simple when viewed from the right viewpoint:

Exercise 1. Given n persons ($n > 0$) and k tasks ($k > 0$).

(a) What is the # of ways to assign a task to each person such that each task has at least 1 person working on it?

(b) What if we additionally want to choose a leader for each task? (The leader should be chosen from the people working on that task.)

(c) What if, instead of choosing leaders, we want to choose a vertical hierarchy for each task? (A vertical hierarchy means a ranking of all people working on this task, from highest to lowest; ties are not allowed.)

Example: Assume that $n = 8$ and $k = 3$. Let our 8 people be $1, 2, 3, 4, 5, 6, 7, 8$, and our 3 tasks be A, B, C.

(a) One option is

task	people working on it
A	$1, 2, 5$
B	3
C	$4, 6, 7, 8$

.

(b) One option is

task	people working on it
A	$1, 2, 5$ with leader 2
B	3 with leader 3
C	$4, 6, 7, 8$ with leader 7

.

(c) One option is

task	people working on it
A	$1 > 5 > 2$
B	3
C	$7 > 8 > 4 > 6$

(where the "$>$" sign means "is ranked above").

For $n = 3$ and $k = 2$, the answers to **(a)**, **(b)** and **(c)** are 6, 12 and 12, respectively.

We will answer this question in Lecture 24.

3. The twelvefold way

So far, we have been answering counting questions one by one, guided mostly by the kinds of numbers that appear in their answers. Let us now try to systematically address a wide class of counting problems. Specifically, we try to solve all reasonable counting problems of the form "how many ways are there to put n balls into k boxes". There are many ways to make this question precise, and most of them lead to interesting problems. The 12 most basic interpretations are usually classified in a table, called the **twelvefold way**.

3.1. What is the twelvefold way?

Convention 3.1.1. For the rest of this chapter, we fix two finite sets A and X. We shall refer to the elements of A as **balls** and to the elements of X as **boxes**.

A **placement** means a way to distribute all balls in A into boxes in X. In other words, it is just a map from A to X. More precisely, this is what we will call an $L \to L$ **placement**, or a **placement of labelled balls into labelled boxes**.

We see immediately that the # of $L \to L$ placements is $|X|^{|A|}$.

For example, the eight $L \to L$ placements of 3 balls $1, 2, 3$ into two boxes $1, 2$

are

$$\begin{aligned}&[123]\ [\ ],\\&[12]\ [3],\\&[13]\ [2],\\&[23]\ [1],\\&[1]\ [23],\\&[2]\ [13],\\&[3]\ [12],\\&[\ ]\ [123].\end{aligned}$$

Here, we are using a semi-visual notation, in which each box is shown as a pair of square brackets, and the numbers between the brackets are the balls placed in this box. The boxes are listed in the obvious order (i.e., the first pair of brackets is box 1, the next is box 2, and so on). Thus, for example, the $L \to L$ placement in which ball 2 lies in box 1 while balls 1 and 3 lie in box 2 is denoted by [2] [13]. (In the 2019 notes, I have used the much more ornate notation $\boxed{\underbrace{\boxed{\quad(2)\quad}}_{\text{box 1}}\quad\underbrace{\boxed{\quad(1)(3)\quad}}_{\text{box 2}}}$ for the same placement; but there is little gained from this flourish. The only problem with the square-brackets notation is that a box [2] could be confused with the set $[2] = \{1, 2\}$; but this confusion is unlikely when there is more than one box.)

Note that the order in which balls are written inside a box doesn't matter; for example, [23] means the same as [32].

Counting $L \to L$ placements is rather boring. But we can vary the problem:

- What if we require our placements (i.e., maps $f : A \to X$) to be injective (i.e., each box contains at most 1 ball) or surjective (i.e., each box contains at least 1 ball)?

- What if the balls are unlabelled (i.e., indistinguishable)?

  In the above example ($A = \{1, 2, 3\}$ and $X = \{1, 2\}$), this means that the three placements

  $$\begin{aligned}&[12]\ [3],\\&[13]\ [2],\\&[23]\ [1]\end{aligned}$$

  no longer count as distinct; instead, they all look the same (namely, they look like "two balls in box 1, and one ball in box 2"). Correspondingly, we now denote them by

  $$[\bullet\bullet]\ [\bullet].$$

- What if the boxes are unlabelled (i.e., indistinguishable)? This means that the two placements

  $$[12]\ [3],$$
  $$[3]\ [12]$$

  no longer count as distinct (since they are the same up to the order of the boxes). I will denote this by underlining the two boxes – i.e., I will write $\underline{[12]\ [3]}$ for both of these placements.

These variants can be combined (e.g., we can assume both balls and boxes to be unlabelled, and we can require the placements to be injective). Thus, we get a total of $3 \cdot 2 \cdot 2 = 12$ different counting problems.[3]

Let us give them names:

- An $L \to L$ **placement** is a placement of labelled balls into labelled boxes.
- An $U \to L$ **placement** is a placement of unlabelled balls into labelled boxes.
- An $L \to U$ **placement** is a placement of labelled balls into unlabelled boxes.
- An $U \to U$ **placement** is a placement of unlabelled balls into unlabelled boxes.

Next time (in Lecture 24), we will define formally what "unlabelled" means, but for now let us give an example:

Example 3.1.2. Let $X = [2]$ and $A = [3]$. Then, let us count how many placements of each kind we have:

	arbitrary	injective	surjective
$L \to L$	8	0	6
$U \to L$	4	0	2
$L \to U$	4	0	3
$U \to U$	2	0	1

In fact:

[3]Yes, we just counted counting problems.

- The $U \to L$ placements are

$$[\bullet\bullet\bullet]\ [\ ],$$
$$[\bullet\bullet]\ [\bullet],$$
$$[\bullet]\ [\bullet\bullet],$$
$$[\ ]\ [\bullet\bullet\bullet].$$

  The second and the third of these are surjective, whereas the first and the fourth are not. None of these placements is injective.

- The $L \to U$ placements are

$$\underline{[123]\ [\ ]},$$
$$\underline{[12]\ [3]},$$
$$\underline{[13]\ [2]},$$
$$\underline{[23]\ [1]}.$$

  The second, third and fourth of these are surjective, whereas the first is not. None of these placements is injective.

- The $U \to U$ placements are

$$\underline{[\bullet\bullet\bullet]\ [\ ]},$$
$$\underline{[\bullet\bullet]\ [\bullet]}.$$

  The second of these is surjective, whereas the first is not. None of these placements is injective.

Not all of the 12 questions in the above table have closed-form answers (for general A and X). However, all of them have at least good recursive answers, and on our quest to find them, we will encounter some nice combinatorial structures.

3.2. $L \to L$

We begin with the first row of the twelvefold way: counting $L \to L$ placements. We repeat their definition:

Definition 3.2.1. An $L \to L$ **placement from** A **to** X means a map from A to X. The value of this map at some $a \in A$ is called the **box in which the ball** a **is placed**.

Now, the following three propositions follow immediately from results that we have proved back in §2.4:

Proposition 3.2.2. We have

$$\begin{aligned}&(\text{\# of } L \to L \text{ placements } A \to X)\\&= (\text{\# of maps } A \to X) = |X|^{|A|}.\end{aligned}$$

Proposition 3.2.3. We have

$$\begin{aligned}&(\text{\# of injective } L \to L \text{ placements } A \to X)\\&= (\text{\# of injective maps } A \to X)\\&= |X|^{\underline{|A|}} = |X| \cdot (|X|-1) \cdot (|X|-2) \cdot \cdots \cdot (|X|-|A|+1).\end{aligned}$$

Proposition 3.2.4. We have

$$\begin{aligned}&(\text{\# of surjective } L \to L \text{ placements } A \to X)\\&= (\text{\# of surjective maps } A \to X) = \operatorname{sur}(|A|, |X|).\end{aligned}$$

Now, to continue with the other three rows of the table, we need to first define formally what “unlabelled” things are.