Math 221 Winter 2023, Lecture 18: Enumeration

website: `https://www.cip.ifi.lmu.de/~grinberg/t/23wd`

4. An introduction to enumeration

4.11. Selections

We now come back to a class of problems that we have posed at the start of Chapter 4 (Lecture 12) but haven't fully answered yet: counting the ways to select a bunch of elements from a given set.

To be more specific, these are problems of the following form: Given an n-element set S, how many ways are there to select k elements from S (where n and k are fixed nonnegative integers)?

The words "k elements" in this question are ambiguous, as they allow for several interpretations:

1. Do we want k arbitrary elements or k distinct elements?
2. Does the order of these k elements matter or not? (In other words, would "$1, 2$" and "$2, 1$" count as two different selections?)

In total, these decisions leave you with 4 options, leading to 4 different problems. In this section, we shall address them all.

4.11.1. Unordered selections without repetition (= without replacement)

Let us begin with the case when we want to select k distinct elements, and the order does not matter. This just means selecting a k-element subset of S. We already know how to count these subsets (Theorem 4.7.4 in Lecture 16):

Theorem 4.11.1. Let $n \in \mathbb{N}$, and let k be any number. Let S be an n-element set. Then,

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

In other words, the # of ways to choose k distinct elements from a given n-element set S, if the order does not matter, is $\binom{n}{k}$.

4.11.2. Ordered selections without repetition (= without replacement)

Now, let us consider the case when the order does matter. Thus, we are looking not for subsets, but for k-tuples. But these k-tuples are not arbitrary k-tuples; they are k-tuples of **distinct** elements. We shall call such k-tuples **injective** (in analogy to injective functions):

Definition 4.11.2. Let $k \in \mathbb{N}$. A k-tuple $(i_1, i_2, \ldots, i_k)$ is said to be **injective** if its k entries $i_1, i_2, \ldots, i_k$ are distinct (i.e., if we have $i_a \neq i_b$ for all $a \neq b$).

For example, the 3-tuple $(6, 1, 2)$ is injective, but $(2, 1, 2)$ is not.

Note that injective k-tuples and injective functions are closely related: A function $f : [k] \to S$ (for a set S and a number $k \in \mathbb{N}$) is injective if and only if the k-tuple $(f(1), f(2), \ldots, f(k))$ is injective.

Next, we introduce another convenient notation:

Definition 4.11.3. Let S be any set, and let $k \in \mathbb{N}$. Then, S^k shall mean the Cartesian product

$$\begin{aligned} \underbrace{S \times S \times \cdots \times S}_{k \text{ times}} &= \{(a_1, a_2, \ldots, a_k) \mid a_1, a_2, \ldots, a_k \in S\} \\ &= \{k\text{-tuples whose all entries belong to } S\}. \end{aligned}$$

For example, $\{5, 6\}^3$ is the set

$$\begin{aligned} &\{5,6\} \times \{5,6\} \times \{5,6\} \\ &= \{(5,5,5), (5,5,6), (5,6,5), (5,6,6), (6,5,5), (6,5,6), (6,6,5), (6,6,6)\}. \end{aligned}$$

None of the 3-tuples (i.e., triples) in this set is injective, but it is easy to find an example where injective k-tuples do appear: For instance, the set $\{1, 2, 3, 4\}^3$ contains both injective 3-tuples such as $(1, 4, 3)$ and non-injective 3-tuples such as $(3, 3, 1)$.

Now, we can define rigorously what we are looking for: A way to select k distinct elements from a given set S, if the order matters, is the same as an injective k-tuple in S^k. We shall now count such ways:

Theorem 4.11.4. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let S be an n-element set. Then,

$$\left(\# \text{ of injective } k\text{-tuples in } S^k\right) = n(n-1)(n-2)\cdots(n-k+1).$$

Example 4.11.5. Applying Theorem 4.11.4 to $n = 5$, $k = 3$ and $S = \{1, 2, 3, 4, 5\}$, we find that

$$\left(\# \text{ of injective 3-tuples in } \{1,2,3,4,5\}^3\right) = 5(5-1)(5-2) = 5 \cdot 4 \cdot 3 = 60.$$

And indeed, there are 60 injective 3-tuples in $\{1, 2, 3, 4, 5\}^3$. For example, $(2, 5, 4)$ and $(5, 3, 2)$ are two of them.

Note that the right hand side in Theorem 4.11.4 is precisely the numerator in the definition of the binomial coefficient $\binom{n}{k}$ (Definition 2.4.1 in Lecture 5), and thus can be rewritten as $k! \cdot \binom{n}{k}$ (since $k!$ is the denominator). Thus, the claim of Theorem 4.11.4 can be restated as

$$\left(\text{\# of injective } k\text{-tuples in } S^k\right) = k! \cdot \binom{n}{k}.$$

Now, how do we prove the theorem? Let us first give an informal proof:

Informal proof of Theorem 4.11.4. Let us look at an example (which is representative of the general case): We let $n = 5$ and $k = 3$ and $S = \{a, b, c, d, e\}$. How many injective k-tuples are there in S^k ? In other words (since $k = 3$): How many injective 3-tuples are there in S^3 ?

Such a 3-tuple has the form (x, y, z), where x, y, z are three distinct elements of S. Let us see how such a 3-tuple can be chosen:

1. First, we choose its first entry x. There are 5 options for this, since S has 5 elements (and x can be any of these 5).
2. Then, we choose its second entry y. There are 4 options for it, since y can be any of the 5 elements of S except for x (because the injectivity of (x, y, z) demands y to be distinct from x).
3. Finally, we choose its third entry z. There are 3 options for it, since z can be any of the 5 elements of S except for x and y (because the injectivity of (x, y, z) demands z to be distinct from x and y) and since x and y are already distinct.

Altogether, we have 5 options at the first step, then 4 options at the second step (no matter which option has been chosen at the first step), and finally 3 options at the third step. Altogether, we can therefore choose our 3-tuple in $5 \cdot 4 \cdot 3$ many different ways, because the numbers of options multiply. Here, we have used a counting rule called "**dependent product rule**", which informally says that if we perform a multi-step construction, and we have

- exactly n_1 options in step 1,
- exactly n_2 options in step 2,
- $\ldots$,
- exactly n_k options in step k,

then the entire construction can be performed in $n_1 n_2 \cdots n_k$ many different ways. We shall not formalize this rule (let alone prove it); the reader can find rigorous versions of this rule in [Loehr11, §1.8] and in [Newste22, Theorem 8.1.19]. However, we shall next give a more rigorous proof of Theorem 4.11.4, which uses induction on k instead of this "dependent product rule" (although the underlying idea is the same). $\square$

Rigorous proof of Theorem 4.11.4. Forget that we fixed S and n. We thus must prove the statement

$$P(k) := \left(\begin{array}{c} \text{"for all } n \in \mathbb{N} \text{ and all } n\text{-element sets } S \text{, we have} \\ (\text{\# of injective } k\text{-tuples in } S^k) = n(n-1)(n-2)\cdots(n-k+1)\text{"} \end{array} \right)$$

for each $k \in \mathbb{N}$. We shall prove this by induction on k.

Base case: We must prove that $P(0)$ holds. In other words, we must prove that for all $n \in \mathbb{N}$ and all n-element sets S, we have

$$\left(\text{\# of injective 0-tuples in } S^0\right) = n(n-1)(n-2)\cdots(n-0+1).$$

But this is an easy exercise in understanding emptiness: Let $n \in \mathbb{N}$, and let S be an n-element set. The only 0-tuple in S^0 is $()$, and this 0-tuple is injective. Thus,

$$\left(\text{\# of injective 0-tuples in } S^0\right) = 1.$$

Comparing this with

$$n(n-1)(n-2)\cdots(n-0+1) = (\text{empty product}) = 1,$$

we obtain $\left(\text{\# of injective 0-tuples in } S^0\right) = n(n-1)(n-2)\cdots(n-0+1)$. Thus, $P(0)$ is proved. This completes the base case.

Induction step: Let k be a positive integer. Assume (as the induction hypothesis) that $P(k-1)$ holds. Our goal is to prove $P(k)$.

We have assumed that $P(k-1)$ holds. In other words, for all $n \in \mathbb{N}$ and all n-element sets S, we have

$$\begin{aligned} &\left(\text{\# of injective } (k-1)\text{-tuples in } S^{k-1}\right) \\ &= n(n-1)(n-2)\cdots(n-(k-1)+1). \end{aligned} \tag{1}$$

Now, let us focus on proving $P(k)$. Thus, we fix an $n \in \mathbb{N}$ and an n-element set S. Our goal is then to prove that

$$\left(\text{\# of injective } k\text{-tuples in } S^k\right) \stackrel{?}{=} n(n-1)(n-2)\cdots(n-k+1).$$

(Again, the question mark atop the equality sign reminds us that this is not proved yet.)

Let $s_1, s_2, \ldots, s_n$ be the n elements of S (listed without repetition). Then, any k-tuple in S^k ends[1] with exactly one of $s_1, s_2, \ldots, s_n$. Hence, by the sum rule, we have

$$\begin{aligned}
&\left(\text{\# of injective } k\text{-tuples in } S^k\right) \\
&= \left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_1\right) \\
&\qquad + \left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_2\right) \\
&\qquad + \cdots \\
&\qquad + \left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_n\right) \\
&= \sum_{i=1}^{n} \left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_i\right). \qquad (2)
\end{aligned}$$

Now, we shall compute the addends in this sum.

Fix any $i \in [n]$. An injective k-tuple in S^k that ends with s_i must have the form

$$(\ldots, s_i),$$

where the "$\ldots$" are $k-1$ distinct elements of $S \setminus \{s_i\}$ (not merely of S, but actually of $S \setminus \{s_i\}$, because if any of them was s_i, then our k-tuple would contain the entry s_i twice and thus fail to be injective). In other words, an injective k-tuple in S^k that ends with s_i is an injective $(k-1)$-tuple in $(S \setminus \{s_i\})^{k-1}$ followed by the entry s_i. Thus, we obtain a map

$$\left\{\text{injective } k\text{-tuples in } S^k \text{ that end with } s_i\right\} \to \left\{\text{injective } (k-1)\text{-tuples in } (S \setminus \{s_i\})^{k-1}\right\},$$
$$(\ldots, s_i) \mapsto (\ldots)$$

(which removes the last entry from our k-tuple and leaves the other entries as they are)[2]. Conversely, we have a map

$$\left\{\text{injective } (k-1)\text{-tuples in } (S \setminus \{s_i\})^{k-1}\right\} \to \left\{\text{injective } k\text{-tuples in } S^k \text{ that end with } s_i\right\},$$
$$(\ldots) \mapsto (\ldots, s_i)$$

(which inserts an s_i after the end of a $(k-1)$-tuple; the result is still injective[3])[4]. These two maps are clearly inverses of each other[5], and thus are bijections.

[1]We say that a k-tuple **ends** with a given element b if b is the last entry of this k-tuple. Note that every k-tuple does indeed have a last entry, since k is positive.

[2]For example, if $k = 4$, then this map sends a k-tuple (a, b, c, s_i) to (a, b, c).

[3]*Proof.* We must show that if we insert an s_i after the end of an injective $(k-1)$-tuple in $(S \setminus \{s_i\})^{k-1}$, then the result is still injective.

Indeed, the only way this could fail is if the newly inserted entry s_i would already appear in the original $(k-1)$-tuple. However, this is impossible, since the original $(k-1)$-tuple belongs to $(S \setminus \{s_i\})^{k-1}$ and thus cannot contain the entry s_i.

[4]For example, if $k = 4$, then this map sends a $(k-1)$-tuple (a, b, c) to (a, b, c, s_i).

[5]because a k-tuple that ends with s_i stays unchanged if we replace its last entry with s_i

Hence, the bijection principle yields

$$\begin{aligned}&\left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_i\right)\\&= \left(\text{\# of injective } (k-1)\text{-tuples in } (S \setminus \{s_i\})^{k-1}\right).\end{aligned}$$

However, recall our induction hypothesis (1). We have $|S| = n$ (since S is an n-element set). Since s_i is an element of S, the set $\{s_i\}$ is a subset of S. Thus, the difference rule (Theorem 4.6.7 **(b)** in Lecture 16) yields

$$|S \setminus \{s_i\}| = \underbrace{|S|}_{=n} - \underbrace{|\{s_i\}|}_{=1} = n - 1,$$

so that $S \setminus \{s_i\}$ is an $(n-1)$-element set, and we have $n - 1 = |S \setminus \{s_i\}| \in \mathbb{N}$. Hence, we can apply (1) to $n - 1$ and $S \setminus \{s_i\}$ instead of n and S (because (1) is a "for all $n \in \mathbb{N}$" statement, not just a statement about the specific n that we have fixed right now!). As a result, we obtain

$$\begin{aligned}&\left(\text{\# of injective } (k-1)\text{-tuples in } (S \setminus \{s_i\})^{k-1}\right)\\&= (n-1)\underbrace{((n-1)-1)}_{=n-2}\underbrace{((n-1)-2)}_{=n-3}\cdots\underbrace{((n-1)-(k-1)+1)}_{=n-k+1}\\&= (n-1)(n-2)(n-3)\cdots(n-k+1).\end{aligned}$$

Combining what we have found, we obtain

$$\begin{aligned}&\left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_i\right)\\&= \left(\text{\# of injective } (k-1)\text{-tuples in } (S \setminus \{s_i\})^{k-1}\right)\\&= (n-1)(n-2)(n-3)\cdots(n-k+1).\end{aligned}$$

Now, forget that we fixed i. We have thus proved that

$$\begin{aligned}&\left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_i\right)\\&= (n-1)(n-2)(n-3)\cdots(n-k+1) \qquad (3)\end{aligned}$$

for **every** $i \in [n]$. Therefore, (2) becomes

$$\begin{aligned}
&\left(\text{\# of injective } k\text{-tuples in } S^k\right) \\
&= \sum_{i=1}^{n} \underbrace{\left(\text{\# of injective } k\text{-tuples in } S^k \text{ that end with } s_i\right)}_{\substack{=(n-1)(n-2)(n-3)\cdots(n-k+1) \\ \text{(by (3))}}} \\
&= \sum_{i=1}^{n} (n-1)(n-2)(n-3)\cdots(n-k+1) \\
&= n \cdot (n-1)(n-2)(n-3)\cdots(n-k+1) \\
&\qquad \left(\text{since } \sum_{i=1}^{n} a = na \text{ for any number } a\right) \\
&= n(n-1)(n-2)\cdots(n-k+1).
\end{aligned}$$

Forget that we fixed n and S. We thus have proved that for all $n \in \mathbb{N}$ and all n-element sets S, we have

$$\left(\text{\# of injective } k\text{-tuples in } S^k\right) = n(n-1)(n-2)\cdots(n-k+1).$$

In other words, we have proved $P(k)$. Thus, the induction step is complete, and Theorem 4.11.4 is proved. □

4.11.3. Intermezzo: Listing n elements

Theorem 4.11.4 tells us that if S is an n-element set, then the # of ways to choose k distinct elements from S, if the order matters, is

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

In particular, applying this to $k = n$, we conclude that the # of ways to choose n distinct elements from S, if the order matters, is

$$n(n-1)(n-2)\cdots(n-n+1) = n! \cdot \underbrace{\binom{n}{n}}_{\substack{=1 \\ \text{(by Corollary 2.5.9} \\ \text{in Lecture 6)}}} = n!.$$

Of course, when we are choosing n distinct elements from an n-element set, we are not actually choosing the elements (since all elements have to be chosen[6]); we are only choosing the order in which we list them. So what we have just shown (if somewhat informally) is the following result:

[6]This follows from Theorem 4.6.7 **(c)** in Lecture 16.

Corollary 4.11.6. Let $n \in \mathbb{N}$. Let S be an n-element set. Then, the # of ways to list the n elements of S in some order (that is, the # of n-tuples that contain each element of S exactly once) is $n!$.

Example 4.11.7. Applying Corollary 4.11.6 to $n = 3$ and $S = \{1, 2, 3\}$, we see that the # of ways to list the 3 numbers $1, 2, 3$ in some order (i.e., the # of 3-tuples that contain each of the numbers $1, 2, 3$ exactly once) is $3! = 6$. And indeed, here are these 6 ways:

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

Corollary 4.11.6 is one of the reasons why factorials are ubiquitous in combinatorics. The $n!$ ways to list the n elements of a given n-element set S are sometimes called the "permutations" of S, but this name is more frequently used for the bijective maps from S to S. (The # of the latter maps is also $n!$, and the two concepts are closely related. For details, see [Math222, §1.7.4 in Lecture 13]. See also [Math222, Lectures 26–28] for much more about permutations.)

4.11.4. Ordered selections with repetition (= with replacement)

We have now solved two variants of our "select k out of n" counting question. We have two more variants to go: the ones where the k elements are arbitrary (not necessarily distinct). Again, we have the choice of caring or not caring about their order.

If we care about their order, then we are just counting all k-tuples in S^k. The answer to this question is simple:

Theorem 4.11.8. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let S be an n-element set. Then,

$$\left(\text{\# of all } k\text{-tuples in } S^k\right) = n^k.$$

Proof. The set S is an n-element set; in other words, $|S| = n$. Now,

$$\begin{aligned}
&\left(\text{\# of all } k\text{-tuples in } S^k\right) \\
&= \left|S^k\right| = \left|\underbrace{S \times S \times \cdots \times S}_{k \text{ times}}\right| \qquad \left(\text{since } S^k \text{ is defined to be } \underbrace{S \times S \times \cdots \times S}_{k \text{ times}}\right) \\
&= \underbrace{|S| \cdot |S| \cdot \cdots \cdot |S|}_{k \text{ times}} \qquad \left(\begin{array}{c}\text{by the product rule for } k \text{ sets} \\ \text{(Theorem 4.6.9 in Lecture 16)}\end{array}\right) \\
&= |S|^k = n^k \qquad (\text{since } |S| = n).
\end{aligned}$$

This proves Theorem 4.11.8. □

4.11.5. Unordered selections with repetition (= with replacement)

Now only one question remains: What is the # of ways to choose k arbitrary elements from an n-element set S if we **don't** care about their order?

There are several equivalent ways to rigorously define what this means:

1. We can define the notion of a **multiset**, which is "like a finite set but allowing an element to be contained multiple times". This is done, e.g., in [Math222, §2.9 (Lectures 21–22)] or (in more detail) in [19fco, §2.11]. Then, a selection of k arbitrary elements from a set S, disregarding the order, can be formalized as a size-k multisubset of the set S.

2. Alternatively, we can define the notion of an **unordered k-tuple**, which is "a k-tuple up to reordering its entries". Formally, these unordered k-tuples are defined as the equivalence classes of usual (i.e., ordered) k-tuples with respect to a certain equivalence relation. (See, e.g., [19fco, Example 3.3.24] for the details.) Then, a selection of k arbitrary elements from a set S, disregarding the order, can be formalized as an unordered k-tuple of elements of S.

3. Finally, if we restrict ourselves to the case when $S = [n]$ (which case is sufficient for all practical purposes, since we can otherwise rename the elements of S as $1, 2, \ldots, n$), then the following "low-tech" solution becomes available: We say that a k-tuple $(i_1, i_2, \ldots, i_k) \in S^k$ is **weakly increasing** (aka **sorted in weakly increasing order**) if it satisfies $i_1 \leq i_2 \leq \cdots \leq i_k$. Now, a selection of k arbitrary elements from $S = [n]$, disregarding the order, can be defined as a weakly increasing k-tuple in S^k (because if we don't care about the order of our k elements, then we can just as well sort them in increasing order, and the result of such a sorting operation is clearly unique[7]).

These three definitions yield different objects, but these objects are equivalent, in the sense that there are bijections from each one to each other. In particular, the # of selections of k arbitrary elements from S (without regard for their order) does not depend on which way we define these selections. Thus, when it comes to counting them, we can pick whatever definition we prefer.

Now that all the requisite warnings and disclaimers have been said, we can finally count these selections:

Theorem 4.11.9. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let S be an n-element set. Then,

$$(\text{\# of all ways to select } k \text{ elements from } S \text{ (if order does not matter)}) = \binom{k+n-1}{k}$$

[7] Clearly if you believe in common sense. Not so clearly if you want a formal proof. See, e.g., [19fco, Exercise 2.11.2] for such a proof.

(where our k elements don't have to be distinct).

Example 4.11.10. Applying Theorem 4.11.9 to $n = 5$ and $k = 2$ and $S = [5] = \{1, 2, 3, 4, 5\}$, we obtain

$$\begin{aligned}&(\text{\# of all ways to select 2 elements from } [5] \text{ (if order does not matter)})\\&= \binom{2+5-1}{2} = \binom{6}{2} = 15.\end{aligned}$$

And indeed, here are these 15 ways:

$$\begin{array}{ccccc}(1,1), & (1,2), & (1,3), & (1,4), & (1,5), \\ & (2,2), & (2,3), & (2,4), & (2,5), \\ & & (3,3), & (3,4), & (3,5), \\ & & & (4,4), & (4,5), \\ & & & & (5,5).\end{array}$$

Here, we have represented each of these selections as a weakly increasing k-tuple in S^k (as explained above).

Informal proof of Theorem 4.11.9 (sketched). For the sake of simplicity, we assume that $S = [n]$ (since otherwise, we can rename the n elements of S as $1, 2, \ldots, n$). Then, as we said above, a selection of k arbitrary elements from $S = [n]$ (disregarding the order) can be defined as a weakly increasing k-tuple in S^k. But a weakly increasing k-tuple in S^k must always look as follows:

$$\left(\underbrace{1, 1, \ldots, 1}_{a_1 \text{ many 1's}}, \underbrace{2, 2, \ldots, 2}_{a_2 \text{ many 2's}}, \ldots, \underbrace{n, n, \ldots, n}_{a_n \text{ many } n\text{'s}}\right)$$

for some numbers $a_1, a_2, \ldots, a_n \in \mathbb{N}$ (in particular, each a_i can be 0, which means that i does not appear in our k-tuple) that satisfy $a_1 + a_2 + \cdots + a_n = k$ (because we want a k-tuple). Such a k-tuple is uniquely determined by these numbers $a_1, a_2, \ldots, a_n$, and conversely, any choice of these numbers $a_1, a_2, \ldots, a_n$ leads to a different k-tuple.

Thus, there is a bijection

$$\begin{aligned}&\text{from } \left\{\text{weakly increasing } k\text{-tuples in } S^k\right\}\\&\text{to } \left\{n\text{-tuples } (a_1, a_2, \ldots, a_n) \in \mathbb{N}^n \text{ satisfying } a_1 + a_2 + \cdots + a_n = k\right\}.\end{aligned}$$

Hence, the bijection principle yields

$$\begin{aligned}
&\left(\text{\# of weakly increasing } k\text{-tuples in } S^k\right)\\
&= \left(\text{\# of } n\text{-tuples } (a_1, a_2, \ldots, a_n) \in \mathbb{N}^n \text{ satisfying } a_1 + a_2 + \cdots + a_n = k\right)\\
&= \left(\text{\# of weak compositions of } k \text{ into } n \text{ parts}\right)\\
&\qquad \left(\begin{array}{c}\text{since the } n\text{-tuples } (a_1, a_2, \ldots, a_n) \in \mathbb{N}^n\\ \text{satisfying } a_1 + a_2 + \cdots + a_n = k\\ \text{are precisely the weak compositions of } k \text{ into } n \text{ parts}\end{array}\right)\\
&= \binom{k+n-1}{k} \qquad \left(\begin{array}{c}\text{by Theorem 4.10.6 in Lecture 17,}\\ \text{applied to } k \text{ and } n \text{ instead of } n \text{ and } k\end{array}\right).
\end{aligned}$$

This proves Theorem 4.11.9 (because these weakly increasing k-tuples in S^k are the ways to select k elements from S (if order does not matter)).

For a rigorous proof, see [19fco, Corollary 2.11.3] (but note that the meanings of the letters n and k are switched in [19fco, Corollary 2.11.3]). □

Theorem 4.11.9 is our fifth combinatorial interpretation of binomial coefficients so far! Previously, we have seen that they count subsets (Theorem 4.7.4 in Lecture 16), lacunar subsets (Theorem 4.9.5 in Lecture 17), compositions (Theorem 4.10.3 in Lecture 17) and weak compositions (Theorem 4.10.6 in Lecture 17). This all is not too surprising, since we proved four of these five theorems using the bijection principle (reducing them to previously proved theorems), but it is impressive to see so many counting problems answered by the same family of numbers.

We have now solved all our four selection problems. We now come to a different counting problem.

4.12. Anagrams and multinomial coefficients

4.12.1. Counting anagrams

An **anagram** of a given word w means a word that consists of the same letters as w but possibly in a different order. For example:

- The anagrams of the word "cat" are "act", "atc", "cat", "cta", "tac" and "tca".
- The word "labl" is an anagram of "ball" (and so are several others).

As you see here, we make no distinction between meaningful and meaningless words. (Also, being logically coherent at the expense of common sense, we consider each word w to be an anagram of itself.)

Now, we can take a given word w and ask how many anagrams w has. For instance:

- How many anagrams does the word "cat" have?

  It has six (and we have just listed them above). In fact, we can put the three letters in any order, and there are 6 possible orders (by Corollary 4.11.6).

- How many anagrams does the word "dud" have?

  It has three ("dud", "ddu", "udd"). Note that the answer does not directly follow from Corollary 4.11.6, since two of the three letters are equal.

- How many anagrams does the word "ball" have?

  It has 12 of them: In fact, if the two "l"s were two different letters, then it would have 24 anagrams (again by Corollary 4.11.6), but since the two "l"s are the same, these 24 anagrams merge into pairs of equal words (you get "ball" twice, you get "blal" twice, etc.), so the answer is $\frac{24}{2} = 12$.

  (Not convinced? Good; it's worth to be skeptical about arguments like this. Still, this argument can be made precise and rigorous. See [Loehr11, first proof of Theorem 1.46] for this.)

- How many anagrams does the word "bookkeeper" have?

  Too many to list by brute force, and the "divide by 2" technique from the previous example gets muddled somewhat as there are several equal letters[8].

  Thus, let us try a new strategy. The word "bookkeeper" has 10 letters. Hence, any anagram of it is a 10-letter word as well. Its letters are

  1 "b", 3 "e"s, 2 "k"s, 2 "o"s, 1 "p" and 1 "r".

  In order to choose an anagram of "bookkeeper", we have to distribute all these letters into 10 positions. In other words, we have to choose which position the 1 "b" will occupy, which positions the 3 "e"s will occupy, and so on. Let us do this step by step:

  - We first choose the position of the 1 "b". There are $\binom{10}{1}$ many options for this, since we need to choose a 1-element subset of the set of all 10 positions.
  - We then choose the positions of the 3 "e"s. There are $\binom{9}{3}$ many options for this, since we need to choose a 3-element subset of the set of all 9 positions not already occupied.

[8] Actually, the technique can be salvaged, but this requires some carefulness that I am too lazy for right now. (Once again, see [Loehr11, first proof of Theorem 1.46].)

– We then choose the positions of the 2 "k"s. There are $\dbinom{6}{2}$ many options for this, since we need to choose a 2-element subset of the set of all 6 positions not already occupied.

– We then choose the positions of the 2 "o"s. There are $\dbinom{4}{2}$ many options for this, since we need to choose a 2-element subset of the set of all 4 positions not already occupied.

– We then choose the positions of the 1 "p". There are $\dbinom{2}{1}$ many options for this, since we need to choose a 1-element subset of the set of all 2 positions not already occupied.

– We then choose the positions of the 1 "r". There are $\dbinom{1}{1}$ many options for this, since we need to choose a 1-element subset of the set of all 1 positions not already occupied.

By the dependent product rule (see the informal proof of Theorem 4.11.4 above), the total # of ways to perform this construction is therefore

$$\begin{aligned}
&\binom{10}{1} \cdot \binom{9}{3} \cdot \binom{6}{2} \cdot \binom{4}{2} \cdot \binom{2}{1} \cdot \binom{1}{1} \\
&= \frac{10!}{1! \cdot 9!} \cdot \frac{9!}{3! \cdot 6!} \cdot \frac{6!}{2! \cdot 4!} \cdot \frac{4!}{2! \cdot 2!} \cdot \frac{2!}{1! \cdot 1!} \cdot \frac{1!}{1! \cdot 0!} \\
&\qquad \text{(by the factorial formula (Theorem 2.5.5 in Lecture 6))} \\
&= \frac{10!}{1! \cdot 3! \cdot 2! \cdot 2! \cdot 1! \cdot 1! \cdot 0!} \qquad \text{(by cancellations)} \\
&= \frac{10!}{1! \cdot 3! \cdot 2! \cdot 2! \cdot 1! \cdot 1!} \qquad \text{(since } 0! = 1\text{)} \\
&= 151\,200.
\end{aligned}$$

Thus, the word "bookkeeper" has $151\ 200 = \dfrac{10!}{1! \cdot 3! \cdot 2! \cdot 2! \cdot 1! \cdot 1!}$ anagrams.

- How many anagrams does the word "anteater" have?

  By the same logic as we just used, it has

$$\frac{8!}{2! \cdot 2! \cdot 1! \cdot 1! \cdot 2!} = 5\ 040 \text{ anagrams.}$$

The same argument works in the general case:

Theorem 4.12.1. Let $s_1, s_2, \ldots, s_n$ be n distinct objects, and let $a_1, a_2, \ldots, a_n$ be n nonnegative integers. Then, the # of tuples that consist of

$$\begin{aligned} & a_1 \text{ copies of } s_1, \\ & a_2 \text{ copies of } s_2, \\ & \ldots, \\ & a_n \text{ copies of } s_n \end{aligned}$$

is

$$\frac{(a_1 + a_2 + \cdots + a_n)!}{a_1! \cdot a_2! \cdot \cdots \cdot a_n!} = \prod_{k=1}^{n} \binom{a_k + a_{k+1} + \cdots + a_n}{a_k}.$$

Informal proof (sketched). Follow the same logic as we used for "bookkeeper" above. To construct such a tuple, we

- first choose the positions for the a_1 many s_1's among its entries (there are $\binom{a_1 + a_2 + \cdots + a_n}{a_1}$ many options for this);
- then choose the positions for the a_2 many s_2's among its entries (there are $\binom{a_2 + a_3 + \cdots + a_n}{a_2}$ many options for this);
- then choose the positions for the a_3 many s_3's among its entries (there are $\binom{a_3 + a_4 + \cdots + a_n}{a_3}$ many options for this);
- and so on, until finally choosing the positions for the a_n many s_n's among its entries (there are $\binom{a_n}{a_n}$ many options for this).

By the dependent product rule, the total # of such tuples is therefore

$$\binom{a_1+a_2+\cdots+a_n}{a_1}\binom{a_2+a_3+\cdots+a_n}{a_2}\binom{a_3+a_4+\cdots+a_n}{a_3}\cdots\binom{a_n}{a_n}$$

$$= \prod_{k=1}^{n} \binom{a_k+a_{k+1}+\cdots+a_n}{a_k}$$

$$= \prod_{k=1}^{n} \frac{(a_k+a_{k+1}+\cdots+a_n)!}{a_k!\,((a_k+a_{k+1}+\cdots+a_n)-a_k)!} \qquad \left(\begin{array}{c}\text{by the factorial formula}\\ \text{(Theorem 2.5.5 in Lecture 6)}\end{array}\right)$$

$$= \prod_{k=1}^{n} \frac{(a_k+a_{k+1}+\cdots+a_n)!}{a_k!\,(a_{k+1}+a_{k+2}+\cdots+a_n)!}$$

$$= \frac{\prod\limits_{k=1}^{n}(a_k+a_{k+1}+\cdots+a_n)!}{\left(\prod\limits_{k=1}^{n} a_k!\right)\left(\prod\limits_{k=1}^{n}(a_{k+1}+a_{k+2}+\cdots+a_n)!\right)}$$

$$= \frac{(a_1+a_2+\cdots+a_n)!\cdot(a_2+a_3+\cdots+a_n)!\cdot\cdots\cdot a_n!}{\left(\prod\limits_{k=1}^{n} a_k!\right)((a_2+a_3+\cdots+a_n)!\cdot(a_3+a_4+\cdots+a_n)!\cdot\cdots\cdot a_n!\cdot 0!)}$$

$$= \frac{(a_1+a_2+\cdots+a_n)!}{\left(\prod\limits_{k=1}^{n} a_k!\right)\cdot 0!} \qquad \left(\begin{array}{c}\text{here, we have cancelled factors that appear}\\ \text{both in the numerator and the denominator}\end{array}\right)$$

$$= \frac{(a_1+a_2+\cdots+a_n)!}{\prod\limits_{k=1}^{n} a_k!} \qquad (\text{since } 0! = 1)$$

$$= \frac{(a_1+a_2+\cdots+a_n)!}{a_1!\cdot a_2!\cdot\cdots\cdot a_n!}.$$

This proves Theorem 4.12.1.

(For a rigorous proof, see [19fco, Proposition 2.12.13]. Note that the objects $s_1, s_2, \ldots, s_n$ are required to be $1, 2, \ldots, n$ in [19fco, Proposition 2.12.13], but this makes no serious difference, since we can always rename them at will.) □

Remark 4.12.2. We can now answer the question "how many prime factorizations does a given number have?" from Lecture 12. For example, consider the number $600 = 2^3 \cdot 3 \cdot 5^2$. A prime factorization of 600 is a tuple that consists of three 2's, one 3 and two 5's, in an arbitrary order. Thus, the # of such prime factorizations is $\frac{6!}{3!\cdot 1!\cdot 2!}$ (by Theorem 4.12.1). Similarly, we can proceed for any positive integer instead of 600.

4.12.2. Multinomial coefficients

The number

$$\frac{(a_1 + a_2 + \cdots + a_n)!}{a_1! \cdot a_2! \cdot \cdots \cdot a_n!}$$

in Theorem 4.12.1 has a name: It is called a **multinomial coefficient**. By Theorem 4.12.1, it is an integer (since it counts something), and can be rewritten as $\prod_{k=1}^{n} \binom{a_k + a_{k+1} + \cdots + a_n}{a_k}$. Note that for $n = 2$, it becomes a binomial coefficient:

$$\frac{(a+b)!}{a! \cdot b!} = \binom{a+b}{a}.$$

Multinomial coefficients have some further properties. There is a standard notation for them: Namely, if $a_1, a_2, \ldots, a_n \in \mathbb{N}$ are any nonnegative integers, and if we set $b = a_1 + a_2 + \cdots + a_n$, then the multinomial coefficient

$$\frac{(a_1 + a_2 + \cdots + a_n)!}{a_1! \cdot a_2! \cdot \cdots \cdot a_n!} = \frac{b!}{a_1! \cdot a_2! \cdot \cdots \cdot a_n!}$$

is denoted by

$$\binom{b}{a_1, a_2, \ldots, a_n}.$$

As already mentioned, multinomial coefficients generalize the binomial coefficients that are found in Pascal's triangle: With our new notation, a binomial coefficient $\binom{n}{k}$ with $n \in \mathbb{N}$ and $k \in \{0, 1, \ldots, n\}$ equals the multinomial coefficient $\binom{n}{k,\ n-k}$. Pascal's identity (Theorem 2.5.3 in Lecture 6, at least for $n > 0$ and $k \in \{0, 1, \ldots, n\}$) thus can be rewritten as

$$\binom{b}{a_1, a_2} = \binom{b-1}{a_1 - 1,\ a_2} + \binom{b-1}{a_1,\ a_2 - 1}$$
$$\text{for } b > 0 \text{ and } a_1, a_2 \in \mathbb{N} \text{ with } a_1 + a_2 = b,$$

where we agree to interpret a multinomial coefficient with a negative number at the bottom to mean 0. An analogue of this identity holds for multinomial coefficients with more parameters:

Theorem 4.12.3 (Recurrence of the multinomial coefficients)**.** Let $b \in \mathbb{N}$ and $a_1, a_2, \ldots, a_n \in \mathbb{N}$ be such that $a_1 + a_2 + \cdots + a_n = b > 0$. Then,

$$\binom{b}{a_1, a_2, \ldots, a_n} = \sum_{i=1}^{n} \underbrace{\binom{b-1}{a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_n}}_{\text{This should be interpreted as 0 if } a_i = 0}.$$

Proof. Nice and fairly easy exercise! (See [19fco, Exercise 2.12.6] for a proof.) □

Just like the binomial coefficients $\binom{n}{k}$ with $n \in \mathbb{N}$ and $k \in \{0, 1, \ldots, n\}$ can be arranged into Pascal's triangle, the multinomial coefficients $\binom{b}{a_1, a_2, \ldots, a_n}$ (for a given n) can be arranged into an n-dimensional analogue of Pascal's triangle, called **Pascal's simplex** (or, for $n = 3$, **Pascal's pyramid**). Theorem 4.12.3 then says that each entry in this simplex (except for the 1 at the apex) is the sum of its n adjacent entries just above it.

Multinomial coefficients owe their name to another fundamental property they satisfy: a generalization of the binomial formula, called the **multinomial formula**:

Theorem 4.12.4 (the multinomial formula)**.** Let $x_1, x_2, \ldots, x_n$ be n numbers. Let $b \in \mathbb{N}$. Then,

$$(x_1 + x_2 + \cdots + x_n)^b = \sum_{\substack{(a_1, a_2, \ldots, a_n) \in \mathbb{N}^n; \\ a_1 + a_2 + \cdots + a_n = b}} \binom{b}{a_1, a_2, \ldots, a_n} x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}.$$

Proof. See [19fco, Theorem 2.12.17] (which gives two references). Here is the simplest proof in a nutshell:

We expand $(x_1 + x_2 + \cdots + x_n)^b$ and collect equal terms. For instance, if $n = 2$ and $b = 3$, then

$$\begin{aligned}
&(x_1 + x_2 + \cdots + x_n)^b \\
&= (x_1 + x_2)^3 \\
&= (x_1 + x_2)(x_1 + x_2)(x_1 + x_2) \\
&= x_1x_1x_1 + x_1x_1x_2 + x_1x_2x_1 + x_1x_2x_2 + x_2x_1x_1 + x_2x_1x_2 + x_2x_2x_1 + x_2x_2x_2 \\
&= x_1^3 + 3x_1^2x_2 + 3x_1x_2^2 + x_2^3.
\end{aligned}$$

What terms do we get for general n and b ? Well, if we expand the product

$$\begin{aligned}
&(x_1 + x_2 + \cdots + x_n)^b \\
&= \underbrace{(x_1 + x_2 + \cdots + x_n)(x_1 + x_2 + \cdots + x_n) \cdots (x_1 + x_2 + \cdots + x_n)}_{b \text{ times}},
\end{aligned}$$

then we obtain the sum of all n^b possible products of the form

$$x_{i_1} x_{i_2} \cdots x_{i_b} \text{ with } i_1, i_2, \ldots, i_b \in [n].$$

Each such product can be rewritten as the monomial $x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, where a_1 is the # of 1's in the b-tuple $(i_1, i_2, \ldots, i_b)$, where a_2 is the # of 2's in this b-tuple, and so on. Moreover, this monomial satisfies $a_1 + a_2 + \cdots + a_n = b$, since the total # of entries of the b-tuple $(i_1, i_2, \ldots, i_b)$ is b.

Thus, expanding $(x_1 + x_2 + \cdots + x_n)^b$, we obtain a sum of monomials of the form $x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ with $a_1 + a_2 + \cdots + a_n = b$, but each such monomial can appear several times in this sum. The total # of copies of a given monomial $x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ that appear in this sum equals the # of all b-tuples that consist of

$$\begin{aligned} &a_1 \text{ copies of } 1, \\ &a_2 \text{ copies of } 2, \\ &\ldots, \\ &a_n \text{ copies of } n \end{aligned}$$

(because of the previous paragraph). But this latter # equals

$$\begin{aligned} &\frac{(a_1 + a_2 + \cdots + a_n)!}{a_1! \cdot a_2! \cdot \cdots \cdot a_n!} && \text{(by Theorem 4.12.1)} \\ &= \frac{b!}{a_1! \cdot a_2! \cdot \cdots \cdot a_n!} && \left(\begin{array}{c} \text{since } a_1 + a_2 + \cdots + a_n = b \text{ (because} \\ \text{our } b\text{-tuple } (i_1, i_2, \ldots, i_b) \text{ has } b \text{ entries in total)} \end{array} \right) \\ &= \binom{b}{a_1, a_2, \ldots, a_n} && \left(\text{by the definition of } \binom{b}{a_1, a_2, \ldots, a_n} \right). \end{aligned}$$

Thus, each monomial $x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ with $a_1 + a_2 + \cdots + a_n = b$ appears exactly $\binom{b}{a_1, a_2, \ldots, a_n}$ times in the sum that we obtain by expanding $(x_1 + x_2 + \cdots + x_n)^b$. Collecting all copies of each monomial in this expansion, we thus obtain

$$(x_1 + x_2 + \cdots + x_n)^b = \sum_{\substack{(a_1, a_2, \ldots, a_n) \in \mathbb{N}^n; \\ a_1 + a_2 + \cdots + a_n = b}} \binom{b}{a_1, a_2, \ldots, a_n} x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}.$$

This proves Theorem 4.12.4. □

We note that this yields a new proof of the binomial formula (Theorem 2.6.1 in Lecture 6), since the latter formula is the particular case of Theorem 4.12.4 for $n = 2$.

Remark 4.12.5. We note that Theorem 4.12.3 can be used to give a second proof of Theorem 4.12.1. Here is a rough outline of this proof:

A tuple that consists of

$$\begin{aligned} &a_1 \text{ copies of } s_1, \\ &a_2 \text{ copies of } s_2, \\ &\ldots, \\ &a_n \text{ copies of } s_n \end{aligned}$$

will be called an $\begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$**-tuple**. Thus, Theorem 4.12.3 is claiming that the # of $\begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$-tuples is $\dbinom{b}{a_1, a_2, \ldots, a_n}$, where $b := a_1 + a_2 + \cdots + a_n$. We shall now prove this by induction on b. The base case ($b = 0$) is trivial (since $b = 0$ entails $a_1 = a_2 = \cdots = a_n = 0$, so we are counting 0-tuples). In the induction step (from $b - 1$ to b), we separate the $\begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$-tuples according to their last entry (just as in our above rigorous proof of Theorem 4.11.4). This last entry is either s_1 or s_2 or $\cdots$ or s_n. Hence, the sum rule yields

$$\begin{aligned}
&\left(\text{\# of } \begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \text{-tuples}\right) \\
&= \sum_{i=1}^{n} \underbrace{\left(\text{\# of } \begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \text{-tuples that end with } s_i\right)}_{\substack{= \left(\text{\# of } \begin{pmatrix} s_1 & s_2 & \cdots & s_{i-1} & s_i & s_{i+1} & \cdots & s_n \\ a_1 & a_2 & \cdots & a_{i-1} & a_i - 1 & a_{i+1} & \cdots & a_n \end{pmatrix} \text{-tuples}\right) \\ \text{(by a bijection argument, just as in the proof of Theorem 4.11.4,} \\ \text{using the bijection that removes the last entry from a tuple)}}} \\
&= \sum_{i=1}^{n} \underbrace{\left(\text{\# of } \begin{pmatrix} s_1 & s_2 & \cdots & s_{i-1} & s_i & s_{i+1} & \cdots & s_n \\ a_1 & a_2 & \cdots & a_{i-1} & a_i - 1 & a_{i+1} & \cdots & a_n \end{pmatrix} \text{-tuples}\right)}_{\substack{= \dbinom{b-1}{a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_n} \\ \text{(by the induction hypothesis if } a_i > 0 \text{, and for obvious reasons if } a_i = 0)}} \\
&= \sum_{i=1}^{n} \dbinom{b-1}{a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_n} \\
&= \dbinom{b}{a_1, a_2, \ldots, a_n} \qquad \text{(by Theorem 4.12.3)},
\end{aligned}$$

which completes the induction step. This proof is less conceptual than the proof we sketched above, but it is easier to formalize, since it does not use the dependent product rule.

References

[19fco] Darij Grinberg, *Enumerative Combinatorics: class notes (Drexel Fall 2019 Math 222 notes)*, 11 September 2022.
http://www.cip.ifi.lmu.de/~grinberg/t/19fco/n/n.pdf

[Loehr11] Nicholas A. Loehr, *Bijective Combinatorics*, Chapman & Hall/CRC 2011.

[Math222] Darij Grinberg, *Math 222: Enumerative Combinatorics, Fall 2022*.
https://www.cip.ifi.lmu.de/~grinberg/t/22fco/

[Newste22] Clive Newstead, *An Infinite Descent into Pure Mathematics*, version 1.0 preview, 26 December 2022.
`https://infinitedescent.xyz`