# Math 222 Fall 2022, Lecture 17: Binomial coefficients

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

# 2. Binomial coefficients (cont'd)

## 2.4. Counting maps (cont'd)

### 2.4.5. Surjective maps (cont'd)

Last time, we introduced the notation $\operatorname{sur}(m,n)$ for the # of surjections[1] from $[m]$ to $[n]$. We noticed that this is also the # of surjective maps from any given $m$-element set to any given $n$-element set.

Let us now compute this number. We shall do this in two different ways, obtaining two different recursive formulas.

*1st approach:* Fix $m \in \mathbb{N}$ and a positive integer $n > 0$. Thus, $n \in [n]$.

Given a surjective map $f : [m] \to [n]$, we let $J_f$ be the set of all $i \in [m]$ such that $f(i) = n$. This $J_f$ is a nonempty subset of $[m]$ (nonempty because $f$ is surjective). Thus, by the sum rule, we have

$$\begin{aligned}&(\text{\# of all surjective maps } f : [m] \to [n]) \\ &= \sum_{\substack{J \subseteq [m];\\ J \neq \varnothing}} \left(\text{\# of all surjective maps } f : [m] \to [n] \text{ such that } J_f = J\right).\end{aligned}$$

Let us now compute the addends in the sum on the RHS.

Fix a nonempty subset $J$ of $[m]$. What is the # of all surjective maps $f : [m] \to [n]$ such that $J_f = J$ ? Imagine trying to construct such a map $f$. Its values $f(j)$ on all elements $j \in J$ are already pre-determined, since each $j \in J$ has to satisfy $f(j) = n$ (since $j \in J = J_f$). Its remaining values (i.e., its values $f(j)$ with $j \notin J$) must belong to $[n-1]$ (indeed, they cannot be $n$, since $j \notin J = J_f$). Thus, $f$ has to send each element of $[m] \setminus J$ to some element of $[n-1]$.

Hence, picking a map $f : [m] \to [n]$ such that $J_f = J$ is tantamount to picking a map from $[m] \setminus J$ to $[n-1]$. Moreover, if we want $f$ to be surjective, we need to ensure that the latter map from $[m] \setminus J$ to $[n-1]$ is surjective (since any element of $[n-1]$ that it does not take as a value will not be a value of $f$ either).

Therefore, in order to construct a surjective map $f : [m] \to [n]$ such that $J_f = J$, all we need to do is to pick a surjective map from $[m] \setminus J$ to $[n-1]$ (and then to set $f(j) = n$ for all $j \in J$). The # of ways to do this is $\operatorname{sur}(m - |J|,\ n-1)$, since $[m] \setminus J$ is an $(m - |J|)$-element set.

[1]"**Surjection**" means "surjective map".

Thus, we obtain

$$\left(\text{\# of all surjective maps } f : [m] \to [n] \text{ such that } J_f = J\right) \\ = \operatorname{sur}\left(m - |J|,\ n-1\right). \tag{1}$$

Forget that we fixed $J$. We have proved this equality (1) for every nonempty subset $J$ of $[m]$. Now let's pick up the above computation again:

$$\begin{aligned}
&(\text{\# of all surjective maps } f : [m] \to [n]) \\
&= \sum_{\substack{J \subseteq [m];\\ J \neq \varnothing}} \underbrace{\left(\text{\# of all surjective maps } f : [m] \to [n] \text{ such that } J_f = J\right)}_{\substack{= \operatorname{sur}(m-|J|,\ n-1) \\ \text{(by (1))}}} \\
&= \sum_{\substack{J \subseteq [m];\\ J \neq \varnothing}} \operatorname{sur}\left(m - |J|,\ n-1\right) \\
&= \sum_{k=1}^{m} \sum_{\substack{J \subseteq [m];\\ J \neq \varnothing;\\ |J| = k}} \operatorname{sur}\Bigg(m - \underbrace{|J|}_{=k},\ n-1\Bigg) \qquad \left(\begin{array}{c}\text{here, we have split up our sum}\\ \text{by the value } |J|\text{, which is}\\ \text{always } \in \{1,2,\ldots,m\}\\ \text{because } J \text{ is nonempty}\end{array}\right) \\
&= \sum_{k=1}^{m} \underbrace{\sum_{\substack{J \subseteq [m];\\ J \neq \varnothing;\\ |J| = k}} \operatorname{sur}\left(m - k,\ n-1\right)}_{= (\text{\# of all nonempty } J \subseteq [m] \text{ satisfying } |J| = k) \cdot \operatorname{sur}(m-k,\ n-1)} \\
&= \sum_{k=1}^{m} \underbrace{\left(\text{\# of all nonempty } J \subseteq [m] \text{ satisfying } |J| = k\right)}_{\substack{= (\text{\# of all } J \subseteq [m] \text{ satisfying } |J| = k) \\ \text{(since } k>0 \text{ ensures that any } k\text{-element set} \\ \text{is automatically nonempty)}}} \cdot \operatorname{sur}\left(m-k,\ n-1\right) \\
&= \sum_{k=1}^{m} \underbrace{\left(\text{\# of all } J \subseteq [m] \text{ satisfying } |J| = k\right)}_{\substack{= (\text{\# of all } k\text{-element subsets of } [m]) = \binom{m}{k} \\ \text{(by the combinatorial interpretation of BCs)}}} \cdot \operatorname{sur}\left(m-k,\ n-1\right) \\
&= \sum_{k=1}^{m} \binom{m}{k} \operatorname{sur}\left(m-k,\ n-1\right).
\end{aligned}$$

Since the LHS of this equality is sur $(m,n)$, we can rewrite this as

$$\begin{aligned}
\operatorname{sur}(m,n) &= \sum_{k=1}^{m} \binom{m}{k} \operatorname{sur}(m-k,\ n-1) \\
&= \sum_{j=0}^{m-1} \underbrace{\binom{m}{m-j}}_{\substack{=\binom{m}{j} \\ \text{(by the symmetry} \\ \text{of Pascal's triangle)}}} \operatorname{sur}(j,\ n-1) \qquad \left(\begin{array}{c}\text{here, we substituted } m-j \\ \text{for } k \text{ in the sum}\end{array}\right) \\
&= \sum_{j=0}^{m-1} \binom{m}{j} \operatorname{sur}(j,\ n-1).
\end{aligned}$$

Let's state this as a proposition:

**Proposition 2.4.11.** Let $m \in \mathbb{N}$, and let $n$ be a positive integer. Then,

$$\operatorname{sur}(m,n) = \sum_{k=1}^{m} \binom{m}{k} \operatorname{sur}(m-k,\ n-1) = \sum_{j=0}^{m-1} \binom{m}{j} \operatorname{sur}(j,\ n-1).$$

This already gives a reasonably fast recursive way for computing sur $(m,n)$. But we can do better, by taking a different approach.

*2nd approach:* Fix two positive integers $m$ and $n$. Let us classify the surjections (= surjective maps) $f : [m] \to [n]$ according to the value $f(m)$.

A surjection $f : [m] \to [n]$ will be called

- **red** if $f(m) = f(i)$ for some $i \in [m-1]$;
- **green** if it is not red (i.e., if $f(m) \neq f(i)$ for all $i \in [m-1]$).

For instance, the surjection $f : [4] \to [3]$ whose one-line notation[2] is $(2,1,3,2)$ is red (since $f(4) = 2$ is also $f(1)$), but the surjection $f : [4] \to [3]$ whose one-line notation is $(2,1,2,3)$ is not (since $f(4) = 3$ is neither $f(1)$ nor $f(2)$ nor $f(3)$).

Here is another way to think of "red" and "green": If we restrict a surjection $f : [m] \to [n]$ to the subset $[m-1]$ (that is, if we remove the value $f(m)$), then we obtain a map from $[m-1]$ to $[n]$ which may or may not be surjective. If it is, then $f$ is red; if it is not, then $f$ is green. In other words, $f$ is green if and only if the value $f(m)$ is "load-bearing" for the surjectivity of $f$. Note that if the map

[2]Recall that the one-line notation of a map $f : [4] \to [3]$ is the 4-tuple $(f(1), f(2), f(3), f(4))$.

$f$ is green, then its restriction to $[m-1]$ will not be a surjection from $[m-1]$ to $[n]$, but it will be a surjection from $[m]$ to $[n] \setminus \{f(m)\}$ (since all elements of $[n]$ except for $f(m)$ will still be taken as values).

Each surjection $f : [m] \to [n]$ is either red or green (but not both). Hence, by the sum rule, we have

$$\begin{aligned} &(\text{\# of surjections } f : [m] \to [n]) \\ &= (\text{\# of red surjections } f : [m] \to [n]) + (\text{\# of green surjections } f : [m] \to [n]) . \end{aligned}$$

Time to compute the addends on the RHS.

Here is a way to construct a red surjection $f : [m] \to [n]$:

- First, we choose the value $f(m)$. There are $n$ options for this, since $f(m)$ can be any element of $[n]$.
- Then, we choose the remaining values $f(1), f(2), \ldots, f(m-1)$. This is tantamount to choosing a surjection from $[m-1]$ to $[n]$ (since the map $f$ should be red, so that even with the value $f(m)$ removed it should still be surjective onto $[n]$). Thus, there are $\operatorname{sur}(m-1, n)$ options for this.

By the dependent product rule[3], the total # of possibilities to make these choices is $n \cdot \operatorname{sur}(m-1,\ n)$. Thus,

$$(\text{\# of red surjections } f : [m] \to [n]) = n \cdot \operatorname{sur}(m-1,\ n) .$$

Here is a way to construct a green surjection $f : [m] \to [n]$:

- First, we choose the value $f(m)$. There are $n$ options for this.
- Then, we choose the remaining values $f(1), f(2), \ldots, f(m-1)$. This is tantamount to choosing a surjection from $[m-1]$ to $[n] \setminus \{f(m)\}$ (since $f$ has to be green, so that its value $f(m)$ cannot be taken by $f$ on any input other than $m$, but all the other elements of $[n]$ still appear as values). Thus, there are $\operatorname{sur}(m-1,\ n-1)$ options for this (since $[n] \setminus \{f(m)\}$ is an $(n-1)$-element set).

Applying the dependent product rule again[4], we see that

$$(\text{\# of green surjections } f : [m] \to [n]) = n \cdot \operatorname{sur}(m-1,\ n-1) .$$

[3]or by the usual product rule

[4]This time, we do need the dependent product rule. The usual product rule would not help, since the options for the values $f(1), f(2), \ldots, f(m-1)$ depend on the choice of $f(m)$.

Altogether, we now have

$$\begin{aligned}
&(\text{\# of surjections } f:[m]\to[n]) \\
&= \underbrace{(\text{\# of red surjections } f:[m]\to[n])}_{=n\cdot\operatorname{sur}(m-1,\ n)} + \underbrace{(\text{\# of green surjections } f:[m]\to[n])}_{=n\cdot\operatorname{sur}(m-1,\ n-1)} \\
&= n\cdot\operatorname{sur}(m-1,\ n) + n\cdot\operatorname{sur}(m-1,\ n-1) \\
&= n\cdot(\operatorname{sur}(m-1,\ n) + \operatorname{sur}(m-1,\ n-1)).
\end{aligned}$$

Since the LHS of this is $\operatorname{sur}(m,n)$, we can rewrite this as follows:

**Proposition 2.4.12.** Let $m$ and $n$ be positive integers. Then,

$$\operatorname{sur}(m,n) = n\cdot(\operatorname{sur}(m-1,\ n) + \operatorname{sur}(m-1,\ n-1)).$$

This recursive equation (which is very similar to Pascal's recurrence for BCs, except for the $n$ factor) makes it easy to fill in a table of $\operatorname{sur}(m,n)$ numbers:

| | | | | | | | | | $n=0$ ↙ | | $n=1$ ↙ | | $n=2$ ↙ | | $n=3$ ↙ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m=0\to$ | | | | | | | | 1 | | 0 | | 0 | | 0 | |
| $m=1\to$ | | | | | | | 0 | | 1 | | 0 | | 0 | | 0 |
| $m=2\to$ | | | | | | 0 | | 1 | | 2 | | 0 | | 0 | |
| $m=3\to$ | | | | | 0 | | 1 | | 6 | | 6 | | 0 | | 0 |
| $m=4\to$ | | | | 0 | | 1 | | 14 | | 36 | | 24 | | 0 | |
| $m=5\to$ | | | 0 | | 1 | | 30 | | 150 | | 240 | | 120 | | 0 |
| $m=6\to$ | | 0 | | 1 | | 62 | | 540 | | 1560 | | 1800 | | 720 | |
| $m=7\to$ | 0 | | 1 | | 126 | | 1806 | | 8400 | | 16800 | | 15120 | | 5040 |

We also observe the following:

**Corollary 2.4.13. (a)** We have $\operatorname{sur}(n,n) = n!$ for each $n\in\mathbb{N}$.
**(b)** The integer $\operatorname{sur}(m,n)$ is divisible by $n!$ for each $m,n\in\mathbb{N}$.

*Proof.* Both parts follow easily by induction using Proposition 2.4.12. (See Corollary 2.4.15 and Exercise 2.4.3 in the 2019 notes for details.) □

**Remark 2.4.14.** Let $m,n\in\mathbb{N}$. The number $\dfrac{\operatorname{sur}(m,n)}{n!}$ (which, by Corollary 2.4.13 **(b)**, is an integer) is often denoted by $\left\{\begin{matrix} m \\ n \end{matrix}\right\}$, and is called a **Stirling number of the 2nd kind**. We will eventually learn more about these numbers.

So far we have seen recursive equations for $\operatorname{sur}(m,n)$. What about explicit formulas? Here is the best one:

**Theorem 2.4.15.** Let $m, n \in \mathbb{N}$. Then,

$$\operatorname{sur}(m, n) = \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} j^m.$$

We will prove this quite soon (Lecture 20, Example 1). (See the 2019 notes for an elementary proof using the above recurrences.)

## 2.5. Sums of powers

Recall that I teased you with a formula for $1^m + 2^m + \cdots + n^m$ using these $\operatorname{sur}(m, n)$ numbers a while ago (Lecture 5, Theorem 1.2.7). We shall now prove it. First, a much more fundamental theorem:

**Theorem 2.5.1.** Let $k \in \mathbb{N}$ and $m \in \mathbb{N}$. Then,

$$k^m = \sum_{i=0}^{m} \operatorname{sur}(m, i) \cdot \binom{k}{i}.$$

*Proof.* Double counting.

Specifically, let us double-count the maps $f : [m] \to [k]$.

On the one hand, the # of such maps is clearly $k^m$ (by Theorem 2.4.1 in Lecture 16).

On the other hand, every map $f : [m] \to [k]$ has a well-defined range (i.e., set of values). This range is a subset of $[k]$. Thus, by the sum rule, we have

$$\begin{aligned}
&(\text{\# of maps } f : [m] \to [k]) \\
&= \sum_{I \subseteq [k]} \underbrace{(\text{\# of maps } f : [m] \to [k] \text{ whose range is } I)}_{\substack{=(\text{\# of surjections from } [m] \text{ to } I) \\ \text{(since the maps from } [m] \text{ to } [k] \text{ whose range is } I \\ \text{are the same as the surjections from } [m] \text{ to } I)}} \\
&= \sum_{I \subseteq [k]} \underbrace{(\text{\# of surjections from } [m] \text{ to } I)}_{=\operatorname{sur}(m,|I|)} \\
&= \sum_{I \subseteq [k]} \operatorname{sur}(m, |I|) = \sum_{i=0}^{k} \sum_{\substack{I \subseteq [k]; \\ |I| = i}} \operatorname{sur}\left(m, \underbrace{|I|}_{=i}\right) \\
&\qquad \text{(here, we have split our sum according to the value of } |I|) \\
&= \sum_{i=0}^{k} \underbrace{\sum_{\substack{I \subseteq [k]; \\ |I| = i}} \operatorname{sur}(m, i)}_{\substack{= \binom{k}{i} \operatorname{sur}(m,i) \\ \text{(since the \# of addends in this sum} \\ \text{is the \# of } i\text{-element subsets of } [k], \\ \text{but this \# is known to be } \binom{k}{i})}} = \sum_{i=0}^{k} \binom{k}{i} \operatorname{sur}(m, i) = \sum_{i=0}^{k} \operatorname{sur}(m, i) \cdot \binom{k}{i}.
\end{aligned}$$

Comparing the two results, we obtain

$$k^m = \sum_{i=0}^{k} \operatorname{sur}(m, i) \cdot \binom{k}{i}. \tag{2}$$

This is almost precisely the claim of Theorem 2.5.1, except that the sum ranges over all $i \in \{0, 1, \ldots, k\}$ instead of over all $i \in \{0, 1, \ldots, m\}$. Fortunately, this little difference is easy to bridge: We have $\binom{k}{i} = 0$ whenever $i > k$ (since $k \in \mathbb{N}$). Thus, the sum $\sum_{i=0}^{k} \operatorname{sur}(m, i) \cdot \binom{k}{i}$ does not change if we extend its range from $i \in \{0, 1, \ldots, k\}$ to all $i \in \mathbb{N}$. Hence,

$$\sum_{i=0}^{k} \operatorname{sur}(m, i) \cdot \binom{k}{i} = \sum_{i \in \mathbb{N}} \operatorname{sur}(m, i) \cdot \binom{k}{i}. \tag{3}$$

However, we also recall (from Lecture 16, Proposition 2.4.10 **(f)**) that

$$\operatorname{sur}(m, i) = 0 \qquad \text{whenever } m < i. \tag{4}$$

Thus, the sum $\sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{k}{i}$ does not change if we extend its range from $i \in \{0, 1, \ldots, m\}$ to all $i \in \mathbb{N}$. Hence,

$$\sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{k}{i} = \sum_{i \in \mathbb{N}} \operatorname{sur}(m,i) \cdot \binom{k}{i}.$$

Comparing this with (3), we find

$$\sum_{i=0}^{k} \operatorname{sur}(m,i) \cdot \binom{k}{i} = \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{k}{i}.$$

Thus, we can rewrite (2) as

$$k^m = \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{k}{i}.$$

This proves Theorem 2.5.1. □

Using the polynomial identity trick, we can generalize Theorem 2.5.1, replacing "$k \in \mathbb{N}$" by "$k \in \mathbb{R}$":

**Theorem 2.5.2.** Let $k \in \mathbb{R}$ and $m \in \mathbb{N}$. Then,

$$k^m = \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{k}{i}.$$

*Proof.* Forget that we fixed $k$. Define the following two polynomials (in one indeterminate $X$, with real coefficients):

$$P := X^m \qquad \text{and} \qquad Q := \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{X}{i}.$$

Then, Theorem 2.5.1 says that $P(k) = Q(k)$ for all $k \in \mathbb{N}$. In other words, $P(x) = Q(x)$ for all $x \in \mathbb{N}$. Thus, by the polynomial identity trick (Lecture 14, Corollary 2.2.5), we have $P = Q$, so that $P(k) = Q(k)$ for all $k \in \mathbb{R}$. But this is precisely the claim of Theorem 2.5.2. □

Note, however, that we cannot extend Theorem 2.5.2 to $m \in \mathbb{R}$, already because $k^m$ is not a polynomial in $m$ for fixed $k$ (but also because $\operatorname{sur}(m,i)$ is not defined for $m \notin \mathbb{N}$, and because $m$ appears as a summation bound on the right hand side).

Now, we can prove our formula for sums of powers (Lecture 5, Theorem 1.2.7). Let us first derive a minor variation of it:

**Theorem 2.5.3.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Then,

$$\sum_{k=0}^{n} k^m = \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{n+1}{i+1}.$$

*Proof.* We have

$$\begin{aligned}
\sum_{k=0}^{n} \underbrace{k^m}_{\substack{=\sum_{i=0}^{m} \operatorname{sur}(m,i)\cdot\binom{k}{i} \\ \text{(by Theorem 2.5.1)}}} &= \sum_{k=0}^{n} \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{k}{i} \\
&= \sum_{i=0}^{m} \sum_{k=0}^{n} \operatorname{sur}(m,i) \cdot \binom{k}{i} \qquad \left(\begin{array}{c}\text{here, we interchanged} \\ \text{the summation signs}\end{array}\right) \\
&= \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \underbrace{\sum_{k=0}^{n} \binom{k}{i}}_{\substack{=\binom{0}{i}+\binom{1}{i}+\cdots+\binom{n}{i} \\ =\binom{n+1}{i+1} \\ \text{(by the hockey-stick identity} \\ \text{(Lecture 7, Theorem 1.3.24))}}} \\
&= \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{n+1}{i+1}, \qquad \text{qed.}
\end{aligned}$$

□

When $m > 0$, the sum $\sum_{k=0}^{n} k^m$ can be rewritten as $\sum_{k=1}^{n} k^m = 1^m + 2^m + \cdots + n^m$, so the claim of Theorem 2.5.3 becomes

$$1^m + 2^m + \cdots + n^m = \sum_{i=0}^{m} \operatorname{sur}(m,i) \cdot \binom{n+1}{i+1}.$$

This is precisely the claim of Theorem 1.2.7 in Lecture 5, just with the variable $k$ renamed as $m$.