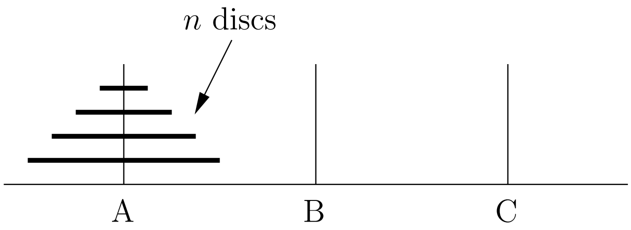
MA241 Combinatorics

Derek Holt
revisions by David Mond and Daan Krammer
September 1, 2006

Contents

1 Introduction: 3 Sample Problems

Example 1: Tower of Hanoi.



Object: Move pile A to B by moving one disc at a time. A disc may never rest on a smaller one. What is the minimal number of moves, T_n?

Strategy: for $n > 1$

1. Move top $(n-1)$ discs $A \to C$ (T_{n-1} moves)
2. Move bottom disk $A \to B$ (1 move)
3. Move top $(n-1)$ $C \to B$ (T_{n-1} moves)

This works and is the quickest solution, as we shall not prove.

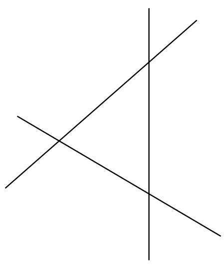
So we have recurrence equations:

$$T_1 = 1\,, \quad T_n = 2T_{n-1} + 1 \tag{1}$$

Technique: Look at some values. Guess the answer. Prove by induction. $T_1 = 1\,,\ T_2 = 3\,,\ T_3 = 7\,,\ T_4 = 15$. Guess $T_n = 2^n - 1$. We can prove this is correct (exercise) using (1).

Example 2: Lines in plane.

What is the maximum number of regions that n infinite lines can divide a plane into? Call this L_n. Get maximum if no lines are parallel and no three lines meet at a point.



$L_0 = 1$ (no line)
$L_1 = 2$
$L_2 = 4$
$L_3 = 7$

What happens when we introduce line n which intersects other $(n-1)$ lines? It gets divided into n sections. Each section divides an old region into two. So

$$L_n = L_{n-1} + n\,. \tag{2}$$

Hence $L_n = L_0 + \sum_{k=1}^{n} k$, for example:

$$\begin{aligned} L_5 &= L_4 + 5 \\ &= L_3 + 4 + 5 \\ &= L_2 + 3 + 4 + 5 \\ &= L_1 + 2 + 3 + 4 + 5 \\ &= L_0 + 1 + 2 + 3 + 4 + 5\,. \end{aligned}$$

Solution:

$$L_n = 1 + \frac{n(n+1)}{2}\,,$$

which can be proved formally by induction.

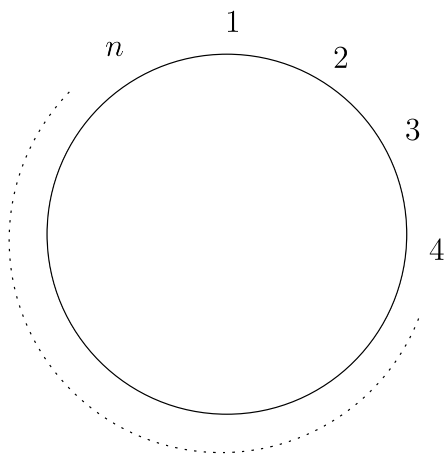
In general, if

$$L_n = L_{n-1} + f(n)\,,$$

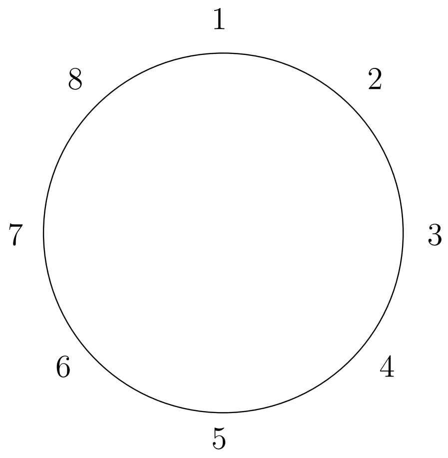
then

$$L_n = L_0 + \sum_{k=1}^{n} f(k)\,.$$

Example 3: Josephus Problem.



n people sit in a circle. Going round clockwise, alternate people are removed. Which person is left last? Call the number of this person J_n. For example,



People who go are: 2, 4, 6, 8, 3, 7, 5. So $J_8 = 1$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
J_n	1	1	3	1	3	5	7	1	3	5	7	9	11	13	15

Guess solution: we can write $n = 2^m + L$ for some $0 \leqslant L < 2^m$. Then $J_n = 2L + 1$.

Case 1: n even, say $n = 2k$.

In the first round all even people are removed. Then k people remain numbered $2i - 1$ $(1 \leqslant i \leqslant k)$. Hence the person left at the end will have number $2J_k - 1$. So $J_{2k} = 2J_k - 1$.

Case 2: n odd, say $n = 2k + 1$ people.

In the first round even people go. Then 1 goes. Then k people are left with numbers $2i + 1$ $(1 \leqslant i \leqslant k)$. $J_{2k+1} = 2J_k + 1$.

The recurrence system to solve is

$$\left.\begin{array}{rcl} J_1 & = & 1 \\ J_{2k} & = & 2J_k - 1 \\ J_{2k+1} & = & 2J_k + 1 \end{array}\right\}$$

We can now prove by induction that $J_n = 2L + 1$.

Case 1: $n = 2k$.

$$n = 2^m + L \Leftarrow k = 2^{m-1} + L/2$$

By inductive hypothesis:

$$J_k = 2(L/2) + 1 = L + 1\,.$$

So

$$J_n = J_{2k} = 2J_k - 1 = 2(L + 1) - 1 = 2L + 1$$

as required.

Case 2: $n = 2k + 1$.

$$n = 2^m + L \Leftarrow k = 2^{m-1} + (L - 1)/2$$

By induction hypothesis:

$$J_k = 2\left(\frac{L-1}{2}\right) + 1 = L\,.$$

So

$$J_n = J_{2k+1} = 2J_k + 1 = 2L + 1\,,$$

as required.

2 Sums

2.1 Notation

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \ldots + a_n \quad \text{and} \quad \sum_{1 \leqslant k \leqslant n} a_k$$

mean the same. We can also write things like:

$$\sum_{\substack{1 \leqslant k \leqslant 12 \\ k \text{ prime}}} \frac{1}{k} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11}.$$

$\sum_k a_k$ means sum over all integers. Usually, only finitely many a_k are non-zero. Otherwise only defined if series converges.

"[Statement]" is defined to be 1 if statement is true and 0 if false. For example:

$$\begin{aligned} [5 \text{ is prime}] &= 1 \\ [2+3=5] &= 1 \\ [2+3=6] &= 0 \end{aligned}$$

$$\sum_{1 \leqslant k \leqslant 12} \frac{[k \text{ is prime}]}{k} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11}$$

$$f : \mathbb{R} \to \mathbb{R}, \quad \text{where} \quad \begin{cases} f(x) = 0, & x \text{ rational} \\ f(x) = 1, & x \text{ irrational} \end{cases}$$

is the same as

$$f(x) = [x \notin \mathbb{Q}].$$

The empty sum is zero. That is, if

$$S_n = \sum_{k=1}^{n} k$$

then

$$S_0 = \text{empty sum} = 0$$

2.2 Sums and Recurrences

Discrete	Continuous
Sum	Integral
Recurrence relation	Differential equation

Sums can be converted to recurrences e.g. $S_n = \sum_{k=1}^{n} 2^{-k}$ is equivalent to $S_0 = 0\,,\ S_n = S_{n-1} + 2^{-n}$. It is useful to go the other way sometimes. This can be done be multiplying by the "summation factor" (which is like the integrating factor in differential equations).

Example 4: (Hanoi)

$$T_0 = 0\,, \quad T_n = 2T_{n-1} + 1$$

Multiply by factor $1/2^n$ to get

$$\frac{T_n}{2^n} = \frac{T_{n-1}}{2^{n-1}} + \frac{1}{2^n}\,.$$

Put $S_n = T_n/2^n$, so

$$S_n \quad = \quad S_{n-1} + 2^{-n}\,,$$

with $S_0 = 0$. So

$$S_n = \sum_{k=1}^{n} 2^{-k} = 1 - \left(\frac{1}{2}\right)^n \ \text{(sum of a G.P.)}$$

So $T_n = 2^n - 1$.

(We used:

$$\sum_{k=1}^{n} ar^k = \frac{a(r^{n+1} - r)}{r - 1}$$

the sum of a G.P. (geometric progression).)

More generally, for the recurrence

$$a_n T_n = b_n T_{n-1} + c_n \quad (b_n \neq 0)\,,$$

we multiply by summation factor

$$\Omega_n = \frac{a_{n-1}a_{n-2}\ldots a_1}{b_n b_{n-1}\ldots b_1}\,,$$

or any constant multiple $c\Omega_n$.Put

$$S_n = \Omega_n a_n T_n$$

to get

$$S_n = S_{n-1} + \Omega_n c_n\,.$$

So

$$S_n = \sum_{k=1}^{n} \Omega_k c_k + S_0\,.$$

Example 5: (from "analysis Quicksort")

$$\begin{aligned} c_0 &= 0 \\ c_n &= n + 1 + \frac{2}{n}\sum_{k=0}^{n-1} c_k \qquad (3) \\ nc_n &= n^2 + n + 2\sum_{k=0}^{n-1} c_k \qquad (4) \end{aligned}$$

Replacing n by $n-1$ this gives

$$(n-1)c_{n-1} = (n-1)^2 + (n-1) + 2\sum_{k=0}^{n-2} c_k \qquad (5)$$

Subtract (5) from (4) to get

$$nc_n = (n+1)c_{n-1} + 2n\,. \qquad (6)$$

This has linear recurrence with $a_n = n\,,\ b_n = (n+1)\,,\ c_n = 2n$. The summation factor is $\Omega_n = \frac{a_{n-1}a_{n-2}\ldots a_1}{b_n b_{n-1}\ldots b_1} = \frac{1}{2n(n+1)}$. So

$$\Omega_n = \frac{1}{n(n+1)}.$$

Multiply (6) by Ω_n to find

$$\frac{c_n}{n+1} = \frac{c_{n-1}}{n} + \frac{2}{n+1}\,.$$

Put

$$S_n = \frac{c_n}{n+1}$$

so

$$S_n = S_{n-1} + \frac{2}{n+1}$$

where $S_0 = 0$. So

$$S_n = \sum_{k=1}^{n} \frac{2}{k+1}\,.$$

Notation: Since there is no formula for this sum, we introduce a new notation. Define the harmonic numbers H_n by

$$H_n = \sum_{k=1}^{n} \frac{1}{k}\,,\ H_0 = 0$$

Note that $H_n \to \infty$ as $n \to \infty$.

Solution to recurrence is

$$\begin{aligned} S_n &= 2(H_{n+1} - 1) \\ c_n &= 2(n+1)(H_{n+1} - 1)\,. \end{aligned}$$

In fact

$$H_n \sim \log_e n + \gamma$$

where

$$\gamma = \text{Euler's constant} \approx 0.5772156$$

2.3 The Perturbation Method

Let $S_n = \sum_{k=0}^{n} a_k$ be a sum. Then

$$S_{n+1} = S_n + a_{n+1} = a_0 + \sum_{k=0}^{n} a_{k+1}$$

The idea is to try and get a formula for the above in terms of S_n. Then solve for S_n.

Example 6: Geometric Series. We define $S_n = \sum_{k=0}^{n} x^k$. Then

$$S_{n+1} = S_n + x^{n+1}.$$

Also

$$S_{n+1} = 1 + \sum_{k=0}^{n} x^{k+1} = 1 + xS_n$$

so

$$S_n = \frac{x^{n+1} - 1}{x - 1} \quad \text{if } x \neq 1$$

Example 7: $T_n = \sum_{k=0}^{n} kx^k$. We have

$$T_{n+1} = T_n + (n+1)x^{n+1}$$

$$\text{and} \quad T_{n+1} = \sum_{k=0}^{n} (k+1)x^{k+1} = xT_n + \sum_{k=0}^{n} x^{k+1} = xT_n + \frac{x^{n+2} - x}{x - 1}$$

$$\text{so} \quad T_n = \frac{-x^{n+2} + x}{(x-1)^2} + \frac{(n+1)x^{n+1}}{(x-1)}.$$

Another method would be to use the fact that

$$T_n = x\frac{d(S_n)}{dx}.$$

2.4 Multiple Sums

Summing over more than one index. For example:

$$\sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k$$

factorises to

$$\left(\sum_{j=1}^{n} a_j\right)\left(\sum_{k=1}^{n} b_k\right).$$

We aim to reduce these to single sums. And more generally, the indices are not independent, for example:

$$\sum_{1\leqslant j<k\leqslant n}(\cdots) = \sum_{j=1}^{n}\sum_{k=j+1}^{n}(\cdots) = \sum_{k=1}^{n}\sum_{j=1}^{k-1}(\cdots).$$

Just like there's Fubini's theorem in calculus (interchanging the order of integration) there is also *interchanging the order of summation*, which is one of the most important tricks in combinatorics. An easy example of this is:

$$\sum_{i=a}^{b}\sum_{j=c}^{d} f(i,j) = \sum_{i=c}^{d}\sum_{j=a}^{b} f(i,j).$$

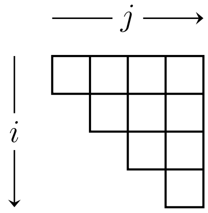
Here's a harder example:

$$\sum_{i=0}^{n}\sum_{j=i}^{n} f(i,j) = \sum_{j=0}^{n}\sum_{i=0}^{j} f(i,j). \tag{7}$$

If you find it hard to find an equation like (7), it can help to draw a *picture* consisting of boxes, one box at position (i,j) for each value of (i,j) involved in the sum. In the case of (7) with $n = 3$ that would be:

(0,0)	(0,1)	(0,2)	(0,3)
	(1,1)	(1,2)	(1,3)
		(2,2)	(2,3)
			(3,3)

=



Example 8:

$$\sum_{1\leqslant j<k\leqslant n} a_j a_k$$

$$\begin{pmatrix} a_1^2 & a_1a_2 & \dots & a_1a_n \\ a_2a_1 & a_2^2 & \dots & a_2a_n \\ \vdots & \vdots & & \vdots \\ a_na_1 & a_na_2 & \dots & a_n^2 \end{pmatrix}$$

We are summing terms above the main diagonal of the matrix. Call this sum $S_{\neg}$. Note that the matrix is symmetric, giving $S_{\neg} = S_{\llcorner}$. So

$$\begin{aligned}\left(\sum_{k=1}^{n} a_k\right)^2 &= \left(\sum_{j=1}^{n} a_j\right)\left(\sum_{k=1}^{n} a_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_j a_k \\ &= S_{\llcorner} + S_{\neg} + \sum_{k=1}^{n} a_k^2 = 2\,S_{\neg} + \sum_{k=1}^{n} a_k^2 .\end{aligned}$$

So

$$S_{\neg} = \frac{1}{2}\left[\left(\sum_{k=1}^{n} a_k\right)^2 - \sum_{k=1}^{n} a_k^2\right].$$

Example 9:

$$\begin{aligned}
\text{If}\quad S_n &= \sum_{1\leqslant j<k\leqslant n} \frac{1}{k-j} \quad \text{then} \\
S_1 &= 0 \text{ (empty sum)} \\
S_3 &= \frac{1}{2-1} + \frac{1}{3-1} + \frac{1}{3-2} = \frac{5}{2}
\end{aligned}$$

$$S_5 = \text{sum of} \quad \begin{array}{c|cccc} & \multicolumn{4}{c}{\longrightarrow k \longrightarrow} \\ \hline \downarrow & 1/1 & 1/2 & 1/3 & 1/4 \\ j & & 1/1 & 1/2 & 1/3 \\ \downarrow & & & 1/1 & 1/2 \\ & & & & 1/1 \end{array} .$$

- Sum by rows:

$$S_n = \sum_{j=1}^{n} \sum_{k=j+1}^{n} \frac{1}{k-j} .$$

  Change of variable: put $k - j = l$.

$$S_n = \sum_{j=1}^{n} \sum_{l=1}^{n-j} \frac{1}{l} = \sum_{j=1}^{n} H_{n-j} = \sum_{j=0}^{n-1} H_j .$$

- Of course, summing by columns gives the same result. However:
- Summing by diagonals:

$$S_n = \sum_{l=1}^{n} \sum_{j=1}^{n-l} \frac{1}{l} = \sum_{j=1}^{n} \frac{(n-j)}{j} = \sum_{j=1}^{n} \left(\frac{n}{j} - 1\right) = nH_n - n .$$

So we have shown that

$$\sum_{j=0}^{n-1} H_j = nH_n - n .$$

Compare this with

$$\int \ln x dx = x \ln x - x + c .$$

2.5 Finite and Infinite Calculus

$$S_1(n) = \sum_{k=0}^{n} k = \frac{n(n+1)}{2} = \frac{n^2}{2} + O(n)$$

$$S_2(n) = \sum_{k=0}^{n} k^2 = \frac{n^3}{3} + O(n^2)$$

In general,

$$S_i(n) = \frac{n^{i+1}}{i+1} + O(n^i)$$

corresponds roughly to

$$\int x^i \, dx = \frac{x^{i+1}}{i+1},$$

but the correspondence is not exact.

Rising and falling powers are easier to sum:

Definition 1: The following notation is not very common, but we will find it quite useful.

- The falling power, $x^{\underline{m}} = \underbrace{x(x-1)\ldots(x-m+1)}_{m \text{ factors}}$ for $x \in \mathbb{R}$.
- The rising power, $x^{\overline{m}} = x(x+1)\ldots(x+m-1)$.

Some properties:

- $x^{\underline{0}} = x^{\overline{0}} = 1$ (empty products are 1).
- for $n \in \mathbb{N}$, $n^{\underline{m}} = \frac{n!}{(n-m)!} = P^n_m = m! \begin{pmatrix} n \\ m \end{pmatrix}$
- $n^{\underline{n}} = n!$
- $n^{\overline{m}} = m! \begin{pmatrix} n+m-1 \\ m \end{pmatrix}$

Then

$$\begin{aligned} (x+1)^{\underline{m}} - x^{\underline{m}} &= x(x-1)\ldots(x-m+2)\Big((x+1)-(x-m+1)\Big) \\ &= m x^{\underline{m-1}} \end{aligned} \tag{8}$$

Similarly

$$x^{\overline{m}} - (x-1)^{\overline{m}} = m x^{\overline{m-1}}. \tag{9}$$

These are similar to differentiation formulae.

Consider

$$\sum_{k=0}^{n} k^{\underline{m}}$$

with

$$k^{\underline{m}} = \frac{(k+1)^{\underline{m+1}} - k^{\underline{m+1}}}{m+1}$$

(by (8) using $(m+1)$ for m).

On summing, most terms cancel. We get

$$\sum_{k=0}^{n} k^{\underline{m}} = \frac{(n+1)^{\underline{m+1}} - 0^{\underline{m+1}}}{m+1} = \frac{(n+1)^{\underline{m+1}}}{m+1}, \quad \text{for } m \geqslant 0\,, \tag{10}$$

which is an exact analogue of integration.

Similarly, summing (9) gives

$$\sum_{k=1}^{n} k^{\overline{m}} = \frac{n^{\overline{m+1}}}{m+1}, \quad \text{for } m \geqslant 0\,. \tag{11}$$

For example, when $m = 2$ and $n = 4$,

$$\begin{aligned} 1.2 + 2.3 + 3.4 + 4.5 &= \frac{6 \times 5 \times 4}{3} \\ &= 40 \end{aligned}$$

We have:

$$\begin{aligned} \sum_{k=1}^{n} k &= \frac{n(n+1)}{2}\,, \\ \sum_{k=1}^{n} k(k+1) &= \frac{n(n+1)(n+2)}{3}\,, \quad \text{etc} \ldots \end{aligned}$$

These can be used to get expressions for $S_m(n)$. For example,

$$S_3(n) = \sum_{k=1}^{n} k^3\,.$$

We must express k^3 in terms of rising (or falling) powers, that is:

$$k^3 = k(k+1)(k+2) - 3k(k+1) + k\,.$$

So

$$\begin{aligned} S_3(n) &= \frac{n(n+1)(n+2)(n+3)}{4} - \frac{3n(n+1)(n+2)}{3} + \frac{n(n+1)}{2} \\ &= \frac{n(n+1)}{4}\left(n^2 + 5n + 6 - 4n - 8 + 2\right) = \frac{n^2(n+1)^2}{4} = (S_1(n))^2. \end{aligned}$$

Note: Since

$$k^{\underline{m}} = m! \begin{pmatrix} k \\ m \end{pmatrix}\,,$$

equation (10) gives

$$\sum_{k=0}^{n} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} n+1 \\ m+1 \end{pmatrix}\,, \quad n \geqslant 0\,.$$

2.6 Negative Exponents in Rising and Falling powers

Motivation:

$$\begin{array}{rcll} x^{\underline{3}} &=& x(x-1)(x-2) & \text{divide by } (x-2) \\ x^{\underline{2}} &=& x(x-1) & \text{divide by } (x-1) \\ x^{\underline{1}} &=& x & \text{divide by } x \\ x^{\underline{0}} &=& 1 & \text{divide by } x+1 \\ x^{\underline{-1}} &=& \frac{1}{x+1} & \dots \\ x^{\underline{-2}} &=& \frac{1}{(x+1)(x+2)} & \dots \end{array}$$

Definition 2:

$$\begin{aligned} x^{-\underline{m}} &= \frac{1}{(x+1)\dots(x+m)} = \frac{1}{(x+1)^{\overline{m}}} \\ x^{-\overline{m}} &= \frac{1}{(x-1)\dots(x-m)} = \frac{1}{(x-1)^{\underline{m}}} \end{aligned}$$

Note: $x^{-\underline{m}} := x^{\underline{-m}}$ for typographical reasons.

It can be checked that equations (8) and (9) are still true for $m \leqslant 0$. So they are true for all $m \in \mathbb{Z}$.

But (10) and (11) become

$$\begin{aligned} (10) \quad \sum_{k=0}^{n} k^{\underline{m}} &= \frac{(n+1)^{\underline{m+1}} - 0^{\underline{m+1}}}{m+1} \\ (11) \quad \sum_{k=1}^{n} k^{\overline{m}} &= \frac{n^{\overline{m+1}} - 0^{\overline{m+1}}}{m+1} \end{aligned}$$

Therefore true for $m \in \mathbb{Z}\,,\ m \neq -1$.

Note: Note that

$$\begin{aligned} 0^{-\overline{m}} &= \frac{1}{(-1)(-2)\dots(-m)} \\ 0^{-\underline{m}} &= \frac{1}{1.2.\dots m} \end{aligned}$$

are non-zero.

Example 10:

$$\sum_{k=0}^{n} k^{-\underline{2}} = \sum_{k=0}^{n} \frac{1}{(k+1)(k+2)} = -\Big(\frac{1}{n+2} - 1\Big) \qquad \text{so} \qquad \sum_{k=0}^{\infty} k^{-\underline{2}} = 1$$

$$\sum_{k=0}^{n} k^{-\underline{3}} = -\frac{1}{2}\Big(\frac{1}{(n+2)(n+3)} - \frac{1}{2}\Big) \qquad \text{so} \qquad \sum_{k=0}^{\infty} k^{-\underline{3}} = \frac{1}{4}$$

What happens for $m = -1$?

$$\sum_{k=0}^{n} k^{\underline{-1}} = \sum_{k=0}^{n} \frac{1}{k+1} = H_{n+1} \sim \log_e(n+1)$$

Compare this with $\int \frac{1}{x}\, dx = \log_e x$.

3 Integer Functions

3.1 Floors and Ceilings

Definition 3: For $x \in \mathbb{R}$:

$$\begin{aligned} \text{Floor}(x) &:= \lfloor x \rfloor = \text{greatest integer } m \leqslant x, \\ \text{Ceiling}(x) &:= \lceil x \rceil = \text{least integer } m \geqslant x. \end{aligned}$$

Example 11:

$$\begin{aligned} \lfloor 3 \rfloor &= \lfloor \pi \rfloor = 3\,, \\ \lceil 3 \rceil &= 3\,, \\ \lceil \pi \rceil &= 4\,, \\ \lfloor -2 \rfloor &= \lfloor -1.4 \rfloor = -2\,, \\ \lceil -1.4 \rceil &= -1\,, \\ \lceil -2 \rceil &= -2\,. \end{aligned}$$

Some easy properties:

$$\lfloor x \rfloor = x \iff x \in \mathbb{Z} \iff \lceil x \rceil = x$$

$$x - 1 < \lfloor x \rfloor \leqslant \lceil x \rceil < x + 1$$

$$\lfloor -x \rfloor = -\lceil x \rceil\,,\ \lceil -x \rceil = -\lfloor x \rfloor$$

$$\lfloor x \rfloor = n \iff n \leqslant x < n + 1 \iff x - 1 < n \leqslant x$$

$$\text{If } n \in \mathbb{Z} \text{ then } \lfloor x + n \rfloor = \lfloor x \rfloor + n.$$

Note that $\lfloor x + y \rfloor$ and $\lfloor x \rfloor + \lfloor y \rfloor$ are not always equal, for example $\lfloor 3/4 + 3/4 \rfloor = 1 \neq 0 = \lfloor 3/4 \rfloor + \lfloor 3/4 \rfloor$.

For $m, n \in \mathbb{Z}$, $n \bmod m$ is defined as the remainder when n is divided by m.

In general, for $x, y \in \mathbb{R}$, $y \neq 0$, we define

$$x \bmod y = x - y\lfloor x/y \rfloor$$

Example 12:

$$\begin{aligned} -5 \bmod 3 &= 1 \quad \text{(not -2)} \\ 5 \bmod -3 &= -1 \quad (x \bmod y \text{ has same sign as } y) \\ -5 \bmod -3 &= -2 \end{aligned}$$

$x \bmod 1 = x - \lfloor x \rfloor =$ fractional part of x. Can define $x \bmod 0 = x$

Exercise. Let $a < b$ be real numbers. We have the four types of intervals

$$\begin{aligned} [a,b] &= \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}, \\ [a,b) &= \{x \in \mathbb{R} \mid a \leqslant x < b\}, \\ (a,b] &= \{x \in \mathbb{R} \mid a < x \leqslant b\}, \\ (a,b) &= \{x \in \mathbb{R} \mid a < x < b\}. \end{aligned}$$

Complete the following table, which expresses the number of integers in such intervals in terms of the floor and ceiling:

$$|[a,b] \cap \mathbb{Z}| \;=\; \qquad\qquad |[a,b) \cap \mathbb{Z}| \;=$$

$$|(a,b] \cap \mathbb{Z}| \;=\; \lfloor b \rfloor - \lfloor a \rfloor, \qquad\qquad |(a,b) \cap \mathbb{Z}| \;=$$

3.2 Floor and Ceiling Problems

Example 13: Is $\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor$ for $x \geqslant 0$? Yes.

Proof:

$$\begin{aligned} m = \lfloor \sqrt{\lfloor x \rfloor} \rfloor &\iff m \leqslant \sqrt{\lfloor x \rfloor} < m+1 \\ &\iff m^2 \leqslant \lfloor x \rfloor < (m+1)^2 \\ &\iff m^2 \leqslant x < (m+1)^2 \\ &\iff m \leqslant \sqrt{x} < (m+1) \\ &\iff m = \lfloor \sqrt{x} \rfloor \end{aligned}$$

Example 14: Is $\lceil \sqrt{\lfloor x \rfloor} \rceil = \lceil \sqrt{x} \rceil$? Not always.

$m+1 = \lceil \sqrt{\lfloor x \rfloor} \rceil \iff m^2 < \lfloor x \rfloor \leqslant (m+1)^2$. If $m^2 \leqslant \lfloor x \rfloor < m^2 + 1$ then $\lfloor x \rfloor = m^2$ and $\sqrt{\lfloor x \rfloor} = m$.

So this equation fails for x satisfying $m^2 < x < m^2 + 1$ for some m, for example $x = 4.3$ or $x = 9.5$.

Example 15: How many integers in the range $1 \leqslant n \leqslant 1000$ satisfy

$$\lfloor \sqrt[3]{n} \rfloor \mid n? \qquad\qquad (*)$$

$a \mid b$ means a divides b exactly where $a, b \in \mathbb{Z}$. Notation: $[5 \ldots 9]$ means $\{5, 6, 7, 8, 9\}$.

For $n \in [1 \ldots 7]$, $\lfloor \sqrt[3]{n} \rfloor = 1$, so $(*)$ holds for all of them.

For $n \in [8 \ldots 26]$, $\lfloor \sqrt[3]{n} \rfloor = 2$, so $(*)$ holds for even n in this range.

For $n \in [27 \ldots 63]$, $\lfloor \sqrt[3]{n} \rfloor = 3$, so $(*)$ holds for multiples of 3 in this range.

$$\vdots$$

For $n \in [729 \ldots 999]$, $\lfloor \sqrt[3]{n} \rfloor = 9$ so $(*)$ holds for multiples of 9 in this range.

For $n = 1000$, $\lfloor \sqrt[3]{n} \rfloor = 10$.

Answer very roughly is

$$7 + \frac{(3^3 - 2^3)}{2} + \frac{(4^3 - 3^3)}{3} + \ldots + \frac{(10^3 - 9^3)}{9} + 1\,.$$

(In fact we want ceilings of terms.)

We can write the problem as that of determining

$$A = \sum_{n=1}^{1000} \left[\lfloor \sqrt[3]{n} \rfloor \mid n \right]$$

Put $k = \lfloor \sqrt[3]{n} \rfloor$

$$\begin{aligned}
A &= 1 + \sum_{k=1}^{9} \sum_{n} [k \mid n][k^3 \leqslant n < (k+1)^3] \\
&= 1 + \sum_{k=1}^{9} \sum_{n,m} [n = mk][k^3 \leqslant n < (k+1)^3] \\
&= 1 + \sum_{k=1}^{9} \sum_{n,m} [n = mk][k^3 \leqslant mk < (k+1)^3] \\
&= 1 + \sum_{k=1}^{9} \sum_{m} [k^2 \leqslant m < k^2 + 3k + 3 + 1/k] \\
&= 1 + \sum_{k=1}^{9} (3k + 4) \\
&= 1 + 3 \times 45 + 36 \\
&= 172
\end{aligned}$$

(Here we have used that $3k+4$ is the number of integers satisfying $k^2 \leqslant m < k^2 + 3k + 3 + 1/k$.)

3.3 Spectra

For $x \in \mathbb{R}$ we define the spectrum, $\operatorname{spec}(x)$, of x to be the set of integers

$$\Big\{ \lfloor x \rfloor\,,\ \lfloor 2x \rfloor\,,\ \lfloor 3x \rfloor\,,\ \lfloor 4x \rfloor\,, \ldots \Big\}$$

Lemma 1 *If* $1 \leqslant x < y$ *then* $\operatorname{spec}(x) \neq \operatorname{spec}(y)$.

Proof
Since $1 \leqslant x$ we have $\lfloor x \rfloor < \lfloor 2x \rfloor < \lfloor 3x \rfloor < \cdots$ and similarly for y. So in

order to prove spec$(x) \neq$ spec(y) it is enough to show that there exists $n \geqslant 0$ with $\lfloor nx \rfloor \neq \lfloor ny \rfloor$.

Since $x < y$, there exists $n \in \mathbb{N}$ with $1/n < y - x$. It follows that $1 < ny - nx$ and therefore $\lfloor ny \rfloor > \lfloor nx \rfloor$ as required. □

Let $\phi = (\sqrt{5}+1)/2$, the golden ratio, about 1.618. Then

$$\begin{aligned} \text{spec}(\phi) &= \{1, 3, 4, 6, 8, 11, 12, 14, \ldots\} \\ \text{spec}(\phi^2) &= \{2, 5, 7, 10, 13, \ldots\} \end{aligned}$$

We notice that

$$\text{spec}(\phi) \cup \text{spec}(\phi^2) = \mathbb{N}$$
$$\text{spec}(\phi) \cap \text{spec}(\phi^2) = \varnothing$$

$\phi + 1 = \phi^2$. Hence $1/\phi + 1/\phi^2 = 1$.

Proposition 2 *Let* $\alpha, \beta \in \mathbb{R} \setminus \mathbb{Q}$ *such that*

$$\frac{1}{\alpha} + \frac{1}{\beta} = 1 \quad \alpha, \beta > 0$$

Then spec$(\alpha) \cup$ spec$(\beta) = \mathbb{N}$ *and* spec$(\alpha) \cap$ spec$(\beta) = \varnothing$

For example, take $\alpha = \sqrt{2}$ and $\beta = 2 + \sqrt{2}$.

$$\begin{aligned} \text{spec}(\sqrt{2}) &= \{1, 2, 4, 5, 7, 8, 9, 11, \ldots\} \\ \text{spec}(2+\sqrt{2}) &= \{3, 6, 10, 13, 17, \ldots\} \end{aligned}$$

Proof
For $\alpha \in \mathbb{R}\,,\ \alpha > 0\,,\ n \in \mathbb{N}$ put

$$N(\alpha, n) = |\{k \in \mathbb{N} \mid \lfloor k\alpha \rfloor \leqslant n\}|\,.$$

We claim that if

$$\frac{1}{\alpha} + \frac{1}{\beta} = 1\,,$$

where α and β are irrational, then

$$N(\alpha, n) + N(\beta, n) = n\,, \quad \forall n \in \mathbb{N}.$$

This will prove the proposition as changing n to $(n+1)$ must increase exactly one of the terms $N(\alpha, n)\,,\ N(\beta, n)$ by 1, that is,

$$\begin{aligned} \text{either } N(\alpha, n+1) &= N(\alpha, n) + 1 \\ \text{or } N(\beta, n+1) &= N(\beta, n) + 1 \end{aligned}$$

but not both. So $n+1$ lies in spec(α) or spec(β) but not both.

Proof of claim:

$$\begin{aligned}
N(\alpha, n) &= \sum_{k>0} [\lfloor k\alpha \rfloor \leqslant n] = \sum_{k>0} [\lfloor k\alpha \rfloor < n+1] \\
&= \sum_{k>0} [k\alpha < n+1] = \sum_{k>0} [0 < k < (n+1)/\alpha] \\
&= \lceil (n+1)/\alpha \rceil - 1.
\end{aligned}$$

(In general the number of integers in $(0, x)$ is $\lceil x \rceil - 1$.)

Since $\alpha \notin \mathbb{Q}$, we have $n + 1/\alpha \notin \mathbb{Z}$, so

$$\lceil (n+1)/\alpha \rceil = (n+1)/\alpha + \eta$$

for some η with $0 < \eta < 1$. Similarly,

$$N(\beta, n) = \lceil (n+1)/\beta \rceil - 1 = (n+1)/\beta + \theta - 1$$

for some $0 < \theta < 1$.

$$\begin{aligned}
N(\alpha, n) + N(\beta, n) &= \frac{n+1}{\alpha} + \eta - 1 + \frac{n+1}{\beta} + \theta - 1 \\
&= n + 1 + \eta + \theta - 1 - 1 \\
&= n + \eta + \theta - 1
\end{aligned}$$

Since this is an integer and $\eta\,,\ \theta \in (0, 1)$, this can only be equal to n. So this completes the proof. □

3.4 Division

Given n cakes to share between k people, how many should each get?

Write $n = kq + r$, where $0 \leqslant r < k$ and $q = \lfloor n/k \rfloor$.

Then r people get $q + 1$ cakes and $k - r$ people get q cakes. Note that

$$\begin{aligned}
\lceil n/k \rceil = \lceil (n-1)/k \rceil = \ldots = \ldots \lceil (n-r+1)/k \rceil = q+1, \\
\lceil (n-r)/k \rceil = \lceil (n-r-1)/k \rceil = \ldots = \lceil (n-k+1)/k \rceil = q.
\end{aligned}$$

So

$$\begin{aligned}
\lceil n/k \rceil + \lceil (n-1)/k \rceil + \ldots + \lceil (n-k+1)/k \rceil &= r(q+1) + (k-r)q \\
&= r + kq = n.
\end{aligned}$$

True for all $n \in \mathbb{Z}\,,\ k \in \mathbb{N}$.

Similarly,

$$\lfloor n/k \rfloor + \lfloor (n+1)/k \rfloor + \ldots + \lfloor (n+k-1)/k \rfloor = n$$

This generalises to $x \in \mathbb{R}$. Replace n by kx to get

$$\lfloor kx \rfloor = \lfloor x \rfloor + \lfloor x + 1/k \rfloor + \lfloor x + 2/k \rfloor + \ldots + \lfloor x + (k-1)/k \rfloor$$

True for all $x \in \mathbb{R}\,,\ k \in \mathbb{N}$.

3.5 Floor and Ceiling Sums

Example 16:

$$\begin{aligned}\sum_{0\leqslant k\leqslant n}\lfloor\sqrt{k}\rfloor &= 0+\underbrace{1+1+1}_{2^2-1^2\text{ terms}}+\underbrace{2+2+2+2+2}_{3^2-2^2\text{ terms}}+3+3+\dots\\ &= \sum j((j+1)^2-j^2)+\text{ some left over}\end{aligned}$$

More precisely, put $a=\lfloor\sqrt{n}\rfloor$. Then

$$\begin{aligned}\sum_{0\leqslant k\leqslant n}\lfloor\sqrt{k}\rfloor &= \sum_{0\leqslant k<a^2}\lfloor\sqrt{k}\rfloor+\underbrace{\sum_{a^2\leqslant k\leqslant n}\lfloor\sqrt{k}\rfloor}_{\text{left over terms}}\\ &= \sum_{0\leqslant j<a} j((j+1)^2-j^2)+(n-a^2+1)a\\ &= \sum_{0\leqslant j<a} j(2j+1)+(n-a^2+1)a\\ &= \sum_{0\leqslant j<a} (2j(j-1)+3j)+a(n-a^2+1)\\ &= \frac{2a(a-1)(a-2)}{3}+\frac{3a(a-1)}{2}+a(n-a^2+1)\\ &= na-\frac{a^3}{3}-\frac{a^2}{2}+\frac{5a}{6}\end{aligned}$$

4 Binomial Coefficients

4.1 Homogeneous trees

Let X be a set. A *permutation* of X is a bijective map $f\colon X \to X$. The set of permutations of X will be written $\mathrm{Bij}(X)$. We will ignore that this is in fact a group (the *symmetric group*).

The factorial numbers $n!$ for $n \geqslant 0$ are defined by

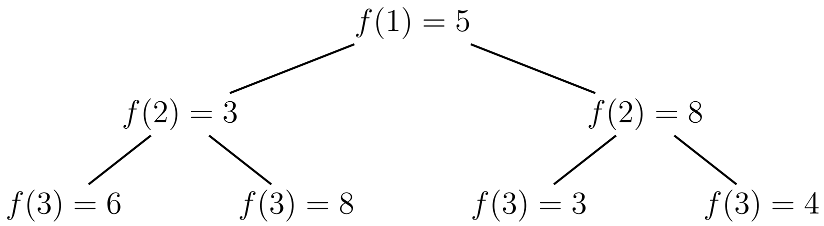
$$n! = 1 \cdot 2 \cdot 3 \cdots n.$$

We have the sets $X_n = \{1, 2, \ldots, n\}$ of n elements.

Proposition 3 *Counting Permutations*

$$|\,\mathrm{Bij}(X_n)| = n!$$

Proof. Imagine we want to choose a permutation $f \in \mathrm{Bij}(X_n)$, and that we do this in n steps. First we choose the value of $f(1)$, then the value of $f(2)$, and so on.



(12)

When choosing $f(1)$, all values of X_n are available, which gives n possibilities.

As a permutation f should satisfy $f(1) \neq f(2)$, the value of $f(2)$ can be chosen from the $n - 1$ possibilities

$$X_n \backslash \{f(1)\}.$$

In general, at the i-th step there are $n + 1 - i$ possible choices

$$X_n \backslash \{f(1), f(2), \ldots, f(i-1)\}.$$

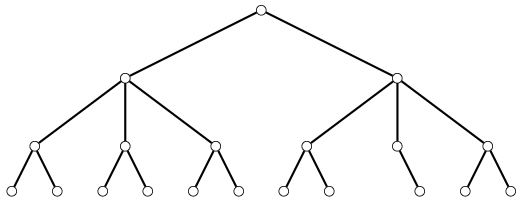
The answer is obtained by *multiplying*[1] these numbers, which gives

$$n \cdot (n-1) \cdot (n-2) \cdots 1 = n!. \qquad \square$$

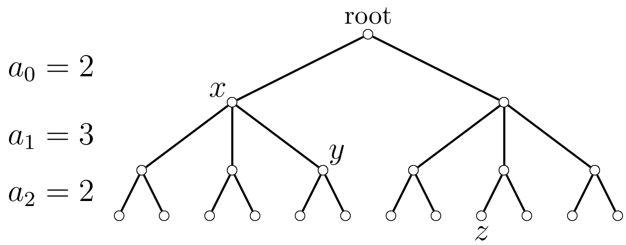
The remarkable thing in this proof is that the number of possible values of $f(2)$ is $n - 1$, regardless the value of $f(1)$ that's just been chosen.

[1]Compare the number of elements of a Cartesian product: $|A \times B| = |A| \times |B|$

An event of choosing an object is a path from top to bottom in a diagram called a *homogeneous (rooted) tree*.



This rooted tree is not homogeneous



This rooted tree is homogeneous

x = vertex,
y = a child of x,
z = a leaf,
number of leaves:
$a_0\, a_1\, a_2 = 2 \cdot 3 \cdot 2 = 12.$

A rooted tree is *homogeneous* if any two *vertices* on the same *level* have the same number of *children.* If a homogeneous tree has n levels, and a vertex on level i has a_i children, then the number of *leaves* (= vertices with no children) is

$$a_0\, a_1\, \cdots a_{n-2}.$$

For nonhomogeneous trees there is no such formula.

So now we understand that (12) is a small part of the homogeneous tree involved in the proof of Proposition 3!

We'll see another homogeneous tree in the following section, and many more after that.

4.2 Binomial Coefficients

Let $r \in \mathbb{R}$, $k \in \mathbb{Z}$. We define the *binomial coefficients* $\binom{r}{k}$ ("r choose k") by

$$\binom{r}{k} = \begin{cases} \dfrac{r(r-1)\dots(r-k+1)}{k!} = \dfrac{r^{\underline{k}}}{k!} & \text{if } k \geqslant 0 \\ 0 & \text{if } k < 0 \end{cases}$$

Proposition 4 *Counting interpretation of binomial coefficients*

Let $n, k \in \mathbb{Z}$, $0 \leqslant k \leqslant n$. The number of k-element subsets of a fixed n-element set is precisely the binomial coefficient $\binom{n}{k}$.

Proof. For the duration of this proof, we write $g(n,k)$ for the number of k-element subsets of $X_n = \{1, \dots, n\}$. So we want to prove $g(n,k) = \binom{n}{k}$.

In Proposition 3 (counting permutations) we saw that $|\operatorname{Bij}(X_n)| = n!$. Let us compute $|\operatorname{Bij}(X_n)|$ another way as follows.

Step 1: Choose $\{f(1), f(2), \ldots, f(k)\}$ (only as a set, not the individual $f(i)$!). By definition, there are $g(n,k)$ ways to do this.

Step 2: Choose the individual values $f(1)$, $f(2)$, $\ldots$, $f(k)$. This boils down to ordering k objects. By Proposition 3 (counting permutations) there are $k!$ ways of doing this. Note that this number does not depend on what happened during step 1: the tree is homogeneous so far!

Step 3: Choose the remaining values $f(k+1)$, $\ldots$, $f(n)$ individually. Again there are $(n-k)!$ ways of doing this.

We have a homogeneous tree, so

$$n! = |\operatorname{Bij}(X_n)| = g(n,k)\, k!\, (n-k)!.$$

and

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

This finishes the proof. □

Note that for all $r \in \mathbb{R}$, $\binom{r}{0} = 1$. Also, $\binom{n}{k} = \binom{n}{n-k}$ if $0 \leqslant k \leqslant n$ are integers.

4.3 Easy Identities

Absorption: for $k \neq 0$,

$$\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}$$

which is clear from the definition.

Proposition 5 *Pascal's Triangle Identity:*

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1} \tag{13}$$

Proof
First assume $n \geqslant k \geqslant 0$, $n \in \mathbb{Z}$. Then

$$\binom{n}{k} = \text{number of ways of choosing } k \text{ things from } n \text{ things}$$

Suppose our set of size n is one red and $(n-1)$ green balls.

$$\binom{n-1}{k-1} = \text{subsets of size } k \text{ which include the red ball}$$

$$\binom{n-1}{k} = \text{subsets of size } k \text{ which do not include the red ball}$$

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

Write (13) as
$$\binom{x}{k} = \binom{x-1}{k} + \binom{x-1}{k-1}.$$

This is a polynomial equation in x of degree k. It is true for all $x \in \mathbb{Z}$ with $x \geqslant 0$, so it has more than k roots. Hence (13) must be an identity. That is, true for all $x \in \mathbb{R}$. □

Using Pascal's triangle identity repeatedly, one easily produces the first few rows of Pascal's triangle:

	0	1	2	3	4	5	6	7	k
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6	4	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
n									$\binom{n}{k}$

Pascal's triangle.

It is not a bad idea to know the first five rows by heart.

Unfold (13), for example:

$$\begin{aligned}\binom{5}{3} &= \binom{4}{3} + \binom{4}{2} \\ &= \binom{4}{3} + \binom{3}{2} + \binom{3}{1} \\ &= \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{1}{0}.\end{aligned}$$

In general (parallel summation),

$$\sum_{k \leqslant n} \binom{r+k}{k} = \binom{r+n+1}{n} \quad \text{for } n \in \mathbb{Z},$$

from summing down diagonal in Pascal's Triangle. Alternatively:

$$\begin{aligned}\binom{5}{3} &= \binom{4}{3} + \binom{4}{2} \\ &= \binom{3}{3} + \binom{3}{2} + \binom{4}{2} \\ &= \binom{2}{2} + \binom{3}{2} + \binom{4}{2}.\end{aligned}$$

In general (upper summation),

$$\sum_{0\leqslant k\leqslant n}\binom{k}{m}=\binom{n+1}{m+1}\quad\text{for } m,n\in\mathbb{Z} \text{ and } m,n\geqslant 0\,,$$

from summation down column in Pascal's Triangle.

Combinatorial explanation of upper summation: choosing $(m+1)$ from $(n+1)$, numbered as $0,1,2\ldots n$. There are $\binom{k}{m}$ ways of making the choice of $(m+1)$ such that the largest number in the chosen set is k. The result follows.

Negation:

$$\begin{aligned} r^{\underline{k}} &= r(r-1)\ldots(r-k+1)\\ &= (-1)^k(-r)(1-r)\ldots(k-1-r)\\ &= (-1)^k(k-1-r)^{\underline{k}} \end{aligned}$$

Dividing by $k!$ gives (upper negation)

$$\binom{r}{k}=(-1)^k\binom{k-1-r}{k}\,. \tag{14}$$

In particular

$$\begin{aligned} \binom{-1}{k} &= (-1)^k \quad \text{for } k\geqslant 0,\\ \binom{-2}{k} &= (-1)^k\binom{k+1}{k}=(-1)^k(k+1). \end{aligned}$$

From this we obtain a formula for the alternating sum of the first m elements in the r'th row of Pascal's triangle:

$$\begin{aligned} \sum_{k\leqslant m}(-1)^k\binom{r}{k} &= \sum_{k\leqslant m}\binom{k-1-r}{k} \quad \text{(upper negation)}\\ &= \binom{-r+m}{m} \quad \text{(parallel summation)}\\ &= (-1)^m\binom{r-1}{m} \quad \text{(upper negation)} \end{aligned}$$

But there is no known simple expression for the same sum without alternating signs,

$$\sum_{k\leqslant m}\binom{r}{k}\,,$$

although of course the special case $m=r$ is obvious :

$$\sum_{k=0}^{r}\binom{r}{k}=2^r\,,\text{the total number of subsets.}$$

Vandermonde Convolution:

$$\sum_{k\in\mathbb{Z}} \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n} \text{ , for } n\in\mathbb{Z}\, r,s\in\mathbb{R}.$$

Combinatorial explanation: suppose we have set of r men, s women. We want to choose n people. Number of subsets with k men and $n-k$ women is $\binom{r}{k}\binom{s}{n-k}$. The result (for the case $r,s\in\mathbb{Z}, r,s\geqslant 0$) follows.

Again it generalises to all $r,s\in\mathbb{R}$ by the roots of polynomial argument.

Theorem 6 *Binomial Theorem.*

$$(x+y)^r = \sum_{k\in\mathbb{Z}} \binom{r}{k} x^k y^{r-k} = y^r \sum_{k\in\mathbb{Z}} \binom{r}{k} \left(\frac{x}{y}\right)^k$$

This is valid if either:

1. $r\in\mathbb{N}$, *series is finite*
2. $r\in\mathbb{R}$ *and* $|x/y|<1$, *series is infinite and converges.*

Another proof of Vandermonde Convolution: Look at coefficient of $x^n y^{r+s-n}$ in

$$\begin{aligned} (x+y)^{r+s} &= (x+y)^r \times (x+y)^s \\ \binom{r+s}{n} &= \sum_k \binom{r}{k}\binom{s}{n-k} \end{aligned}$$

This covers all top 10 identities, except "trinomial revision". Let $m,k\in\mathbb{Z}$, $r\in\mathbb{R}$, $m\geqslant k$.

$$\begin{aligned} \binom{r}{m}\binom{m}{k} &= \frac{r(r-1)\dots(r-m+1)}{m!} \times \frac{m!}{k!(m-k)!} \\ &= \frac{r(r-1)\dots(r-k+1)}{k!} \times \frac{(r-k)\dots(r-m+1)}{(m-k)!} \\ &= \binom{r}{k}\binom{r-k}{m-k} \end{aligned}$$

(if $m<k$ then both sides are 0).

The top ten binomial coefficient identities.

Factorial expansion	$\binom{n}{k}$	$=$	$\dfrac{n!}{k!\,(n-k)!}$,	$n \geqslant k \geqslant 0$ integers
Symmetry	$\binom{n}{k}$	$=$	$\binom{n}{n-k}$,	$n \geqslant k \geqslant 0$ integers
Absorption	$\binom{r}{k}$	$=$	$\dfrac{r}{k}\binom{r-1}{k-1}$,	$k > 0$ integer
Pascal's triangle identity	$\binom{r}{k}$	$=$	$\binom{r-1}{k} + \binom{r-1}{k-1}$,	k integer
Upper negation	$\binom{r}{k}$	$=$	$(-1)^k\binom{k-r-1}{k}$,	k integer
Trinomial revision	$\binom{r}{m}\binom{m}{k}$	$=$	$\binom{r}{k}\binom{r-k}{m-k}$,	m, k integers
Binomial theorem	$(x+y)^r$	$=$	$\sum_k \binom{r}{k} x^k y^{r-k}$,	$r > 0$ integer or $\lvert x/y \rvert < 1$
Parallel summation	$\sum_{k \leqslant n}\binom{r+k}{k}$	$=$	$\binom{r+n+1}{n}$,	n integer
Upper summation	$\sum_{0 \leqslant k \leqslant n}\binom{k}{m}$	$=$	$\binom{n+1}{m+1}$,	$m, n \geqslant 0$ integers
Vandermonde convolution	$\binom{r+s}{n}$	$=$	$\sum_k \binom{r}{k}\binom{s}{n-k}$,	n integer.

4.4 Some more complicated identities

Example 17: Let

$$S = \sum_{k=0}^{m} \binom{m}{k} \Big/ \binom{n}{k}, \text{ for } n \geqslant m \geqslant 0$$

By trinomial revision

$$\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k},$$

So

$$\binom{m}{k} \Big/ \binom{n}{k} = \binom{n-k}{m-k} \Big/ \binom{n}{m}$$

Hence

$$\binom{n}{m} \times S = \sum_{k=0}^{m} \binom{n-k}{m-k}.$$

This looks like parallel summation. We need to change variable in summation. Put $l = m - k$, $k = m - l$.

$$\begin{aligned}
\binom{n}{m} \times S &= \sum_{l=0}^{m} \binom{n-m+l}{l} \\
&= \binom{n-m+m+1}{m} \text{ (parallel summation)} \\
&= \binom{n+1}{m} \\
\Rightarrow S &= \binom{n+1}{m} \Big/ \binom{n}{m} \\
&= \frac{n+1}{n+1-m}
\end{aligned}$$

Example 18:

$$S = \sum_{k=0}^{m} k \binom{n-k-1}{n-m-1} \Big/ \binom{n}{m} \quad \text{for } n > m \geqslant 0\,.$$

If the first k in each summand were only $(n-k)$ we could use absorption. So we can write $k = n - (n-k)$ and then

$$S \binom{n}{m} = S_1 - S_2$$

where

$$S_1 = \sum_{k=0}^{m} n \binom{n-k-1}{n-m-1}$$

and

$$S_2 = \sum_{k=0}^{m} (n-k) \binom{n-k-1}{n-m-1}\,.$$

By absorption,

$$S_2 = \sum_{k=0}^{m} (n-m) \binom{n-k}{n-m}$$

Now we make a change of variable:

$$l = n-k\,, \quad k = n-l\,.$$

$$S_2 = (n-m) \sum_{l=n-m}^{n} \binom{l}{n-m}$$

But if $l < n-m$, then $\binom{l}{n-m} = 0$. So

$$S_2 = (n-m) \sum_{l=0}^{n} \binom{l}{n-m} = (n-m) \binom{n+1}{n-m+1}$$

by upper summation.

$$S_1 = n \sum_{k=0}^{m} \binom{n-k-1}{n-m-1}$$

which is similar to S_2 with $(n-1)$ in place of n. Put

$$n-k-1 = l\,, \quad k = n-l-1$$

$$S_1 = n \sum_{l=0}^{n-1} \binom{l}{n-m-1} = n \binom{n}{n-m} \quad \text{by upper summation}$$

So

$$\begin{aligned} S\binom{n}{m} &= S_1 - S_2 \\ &= n\binom{n}{n-m} - (n-m)\binom{n+1}{n-m+1} \\ &= n\binom{n}{n-m} - \frac{(n+1)(n-m)}{(n-m+1)}\binom{n}{n-m} \\ &= \frac{m}{n-m+1}\binom{n}{n-m} \\ &= \frac{m}{n-m+1}\binom{n}{m}\,. \end{aligned}$$

So

$$S = \frac{m}{n-m+1}\,.$$

4.5 Derangements

Theorem 7 *Inversion Formula.*

Let $f, g : \mathbb{N}_0 \to \mathbb{R}$ be functions (here $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$). Then

$$g(n) = \sum_{k\in\mathbb{Z}} (-1)^k \binom{n}{k} f(k) \quad \text{for all } n \geqslant 0 \qquad \Longleftrightarrow$$

$$f(n) = \sum_{k\in\mathbb{Z}} (-1)^k \binom{n}{k} g(k) \quad \text{for all } n \geqslant 0.$$

Note: The sums are the same as $\sum_{k=0}^{n}$ because $\binom{n}{k} = 0$ otherwise. So they are finite sums.

Proof
By symmetry, enough to prove LHS ⇒ RHS, so assume LHS. Then

$$\begin{aligned}\sum_k \binom{n}{k}(-1)^k g(k) &= \sum_k \binom{n}{k}(-1)^k \sum_j \binom{k}{j}(-1)^j f(j)\\ &= \sum_j f(j)\sum_k (-1)^{j+k}\binom{n}{k}\binom{k}{j}\end{aligned}$$

where we have replaced n by k, and k by j in LHS. Now

$$\binom{n}{k}\binom{k}{j} = \binom{n}{j}\binom{n-j}{k-j},$$

so by trinomial revision the sum is

$$\begin{aligned}&\sum_j f(j)\binom{n}{j}\sum_k (-1)^{j+k}\binom{n-j}{k-j}\\ &= \sum_j f(j)\binom{n}{j}\sum_{l+j}(-1)^{l+2j}\binom{n-j}{l}\end{aligned}$$

where we have replaced k by $l+j$ and $k-j$ by l.

Since j is constant in $\sum_{l+j}$ term, $\sum_{l+j}$ is the same as $\sum_l$ (remember the range of summation is all of $\mathbb{Z}$). So sum is

$$\begin{aligned}&= \sum_j f(j)\binom{n}{j}\sum_l (-1)^l\binom{n-j}{l}\\ &= \sum_j f(j)\binom{n}{j}\sum_{l=0}^{n-j}(-1)^l\binom{n-j}{l}\end{aligned}$$

(since $\binom{n-j}{l} = 0$ for other terms).

But we have shown that

$$\sum_{k=0}^{m}\binom{r}{k}(-1)^k = (-1)^m\binom{r-1}{m}$$

So our sum is

$$\sum_j f(j)\binom{n}{j}\binom{n-j-1}{n-j}. \qquad (*)$$

Note that for $n \in \mathbb{Z}$,

$$\binom{n-1}{n} = \begin{cases} 0 \text{ for } n \neq 0\\ 1 \text{ for } n = 0\end{cases}$$

That is,

$$\binom{n-1}{n} = [n=0]$$

— only non-zero term in $(*)$ is when $j = n$. So

$$(*) = f(n)\binom{n}{n} = f(n)\,, \text{ for all } n \geqslant 0\,.$$

So LHS $\Rightarrow$ RHS as required. $\square$

There are $n!$ permutations of a set of size n. A permutation is called a *derangement* if it does not fix any point in the set.

Let $D(n)$ be the number of derangements. Can we get a formula for $D(n)$?

More generally, let $h(n,r)$ be the number of permutations fixing exactly r points, $0 \leqslant r \leqslant n$.

So $D(n) = h(n,0)$ (note that $h(n,n-1) = 0$).

n	$h(n,0)$	$h(n,1)$	$h(n,2)$	$h(n,3)$	$h(n,4)$
1	0	1			
2	1	0	1		
3	2	3	0	1	
4	9	8	6	0	1

Example 19: For $n = 4$ the permutations $(1,2,3,4)$, $(1,2,4,3)$, $(1,3,4,2)$, $(1,3,2,4)$, $(1,4,2,3)$, $(1,4,3,2)$, $(1,2)(3,4)$, $(1,4)(3,2)$, $(1,3)(2,4)$ fix 0 points. So there are 9 derangements.

$(1,2,3)$, $(1,3,2)$, $(1,2,4)$, $(1,4,2)$, $(1,3,4)$, $(2,3,4)$, $(2,4,3)$, $(1,4,3)$ fix 1 point.

$(1,2)$, $(1,3)$, $(1,4)$, $(2,3)$, $(2,4)$, $(3,4)$ fix 2 points

Identity fixes 4 points

Note that $h(n,r) = \binom{n}{r} D(n-r)$ because we choose r fixed points in $\binom{n}{r}$ ways then derange the other $n-r$ points in $D(n-r)$ ways. Clearly the total number of permutations:

$$\begin{aligned} n! &= \sum_{k=0}^{n} h(n,k) = \sum_{k=0}^{n} \binom{n}{k} D(n-k) \\ &= \sum_{k} \binom{n}{k} D(n-k) = \sum_{k} \binom{n}{n-k} D(n-k) \\ &= \sum_{k} \binom{n}{k} D(k) \text{ replacing } k \text{ with } n-k\,. \end{aligned}$$

We can now use Theorem (7) with $g(n) = n!$ and $f(n) = (-1)^n D(n)$ to conclude that,

$$f(n) = (-1)^n D(n) = \sum_k \left(\begin{array}{c} n \\ k \end{array} \right) (-1)^k k!$$

$$\text{so} \quad D(n) = \sum_{k=0}^{n} \frac{n!}{(n-k)!}(-1)^{n+k} = \sum_{k=0}^{n} \frac{n!}{(n-k)!}(-1)^{n-k}$$

$$= \sum_{k=0}^{n} \frac{n!}{k!}(-1)^k \,, \text{ replacing } n-k \text{ by } k \,.$$

Example 20:

$$D(1) = 0$$

$$D(2) = 2!\Big(1 - \frac{1}{1!} + \frac{1}{2!}\Big) = 1$$

$$D(3) = 3!\Big(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!}\Big) = 2$$

$$D(4) = 4!\Big(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}\Big) = 9$$

Note that $D(n)/n!$ is the proportion of derangements as a subset of all permutations. It tends to $1/e$ rapidly as $n \to \infty$. In fact

$$D(n) = \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor .$$

4.6 Multinomial coefficients

These are generalisations of binomial coefficients. We define

$$\binom{n_1 + \cdots + n_k}{n_1, \ldots, n_k} := \frac{(n_1 + \cdots + n_k)!}{n_1! \cdots n_k!}.$$

Note that

$$\binom{k+\ell}{k, \ell} = \binom{k+\ell}{k},$$

the binomial coefficient.

We've seen before that $\binom{n}{k}$ is the number of k-element subsets of an n-element set. The following generalises this to multinomial coefficients, necessarily with a somewhat different language.

Let $X(m)$ denote a set of m elements, and let $A(n_1, \ldots, n_k)$ denote the set of k-tuples $(Y_1, \ldots, Y_k)$ of disjoint sets whose union is $X(n_1 + \cdots + n_k)$, and such that $|Y_i| = n_i$ for all i.

Proposition 8 *Combinatorial description of multinomial coefficients.*

$$\binom{n_1+\cdots+n_k}{n_1,\ldots,n_k} = |A(n_1,\ldots,n_k)|.$$

Exercise. Prove the above proposition by a homogeneous tree.

The binomial theorem also has its generalisation, which looks as follows.

Theorem 9 *Multinomial theorem.*

Let $k \geqslant 1$ and $m \geqslant 0$ be integers. Then

$$(x_1+\cdots+x_k)^m = \sum \binom{m}{n_1,\ldots,n_k} x_1^{n_1}\cdots x_k^{n_k}$$

where the sum is over all k-tuples $(n_1,\ldots,n_k)$ of nonnegative integers whose sum is m. □

Exercises about multinomial coefficients.

1. Prove
$$\binom{k+\ell+m}{k,\ell,m} = \binom{k+\ell+m}{k+\ell,m}\binom{k+\ell}{k,\ell}$$
and deduce trinomial revision.

2. Prove the multinomial theorem, using the binomial theorem and induction on k.

3. Prove the multinomial theorem, using differentiation and induction on m.

4. Prove:
$$\begin{aligned}\binom{n_1+\cdots+n_k}{n_1,\ldots,n_k} &= \binom{n_1+\cdots+n_k-1}{n_1-1,n_2,\ldots,n_k}+\binom{n_1+\cdots+n_k-1}{n_1,n_2-1,n_3,\ldots,n_k} \\ &\quad+\cdots+\binom{n_1+\cdots+n_k-1}{n_1,\ldots,n_{k-2},n_{k-1}-1,n_k} \\ &\quad+\binom{n_1+\cdots+n_k-1}{n_1,\ldots,n_{k-1},n_k-1} \qquad (15)\end{aligned}$$

5. Prove the multinomial theorem, using induction on m and (15).

6. Fix $n_1,\ldots,n_{k-1} \in \mathbb{Z}_{\geqslant 0}$. Prove that the map
$$n_k \mapsto \binom{n_1+\cdots+n_k}{n_1,\ldots,n_k}$$
is a polynomial. (This tells us how a more general multinomial coefficient should be defined where one variable is real and all other variables are nonnegative integers.)

5 Special Numbers

5.1 Stirling Numbers

Stirling numbers of the first and second kinds are written as

$$\text{first kind} \left[\begin{array}{c} n \\ k \end{array} \right], \quad \text{second kind} \left\{ \begin{array}{c} n \\ k \end{array} \right\}$$

They are defined only for $n \in \mathbb{N}_0$, $k \in \mathbb{Z}$. They are 0 for $k < 0$ and $k > n$. Let $X_n = \{1, 2, 3, \ldots n\}$ be a set of size n.

Stirling numbers of the second kind

Definition 4: A *partition* of a set X is an equivalence relation on that set. The equivalence classes are called *parts*. $\left\{ {n \atop k} \right\}$ is the number of ways of partitioning X_n into k non-empty subsets.

Example 21: $n = 4$, $k = 2$. We have the partitions $\{1, 2, 3\} \cup \{4\}$, $\{1, 2, 4\} \cup \{3\}$, $\{1, 3, 4\} \cup \{2\}$, $\{2, 3, 4\} \cup \{1\}$, $\{1, 2\} \cup \{3, 4\}$, $\{1, 3\} \cup \{2, 4\}$, $\{1, 4\} \cup \{2, 3\}$.

So $\left\{ \begin{array}{c} 4 \\ 2 \end{array} \right\} = 7$.

Properties:

$$\begin{aligned} \left\{ \begin{array}{c} 0 \\ 0 \end{array} \right\} &= 1 \\ \left\{ \begin{array}{c} n \\ 0 \end{array} \right\} &= 0\,, \text{for } n > 0 \\ \left\{ \begin{array}{c} n \\ 1 \end{array} \right\} &= 1\,,\ \text{for } n > 0 \\ \left\{ \begin{array}{c} n \\ 2 \end{array} \right\} &= 2^{n-1} - 1\,,\ \text{for } n \geqslant 2 \\ &= \frac{2^n - 2}{2} \end{aligned}$$

(note that $2^n - 2$ is the number of non empty proper subsets of X_n.)

Proposition 10 *Basic Recurrence Relation:*

$$\left\{ \begin{array}{c} n \\ k \end{array} \right\} = k \left\{ \begin{array}{c} n-1 \\ k \end{array} \right\} + \left\{ \begin{array}{c} n-1 \\ k-1 \end{array} \right\} \tag{16}$$

Proof
Consider a partition of X_n into k non-empty subsets $(n, k > 0)$.

Either

1. $\{n\}$ is one of the subsets. Then there are $\left\{{n-1 \atop k-1}\right\}$ ways of decomposing $X_n \setminus \{n\}$ into $k-1$ non empty subsets.
2. n is in a larger subset. Then we decompose $X_n \setminus \{n\}$ into k non-empty subsets in $\left\{{n-1 \atop k}\right\}$ ways. We can place n in any of these k subsets. So there are $\left\{{n-1 \atop k}\right\}$ ways of doing this.

Result follows. □

n	$\left\{{n \atop 0}\right\}$	$\left\{{n \atop 1}\right\}$	$\left\{{n \atop 2}\right\}$	$\left\{{n \atop 3}\right\}$	$\left\{{n \atop 4}\right\}$	$\left\{{n \atop 5}\right\}$
0	1					
1	0	1				
2	0	1	1			
3	0	1	3	1		
4	0	1	7	6	1	
5	0	1	15	25	10	1

Example 22: Unfolding is the analogue of upper and parallel summation. Unfolding (16) ℓ times ($\ell \leqslant k$) yields

$$\begin{aligned}\left\{{n \atop k}\right\} &= k\left\{{n-1 \atop k}\right\} + \left\{{n-1 \atop k-1}\right\} \\ &= k\left\{{n-1 \atop k}\right\} + (k-1)\left\{{n-2 \atop k-1}\right\} + \left\{{n-2 \atop k-2}\right\} = \\ &= \Big(\sum_{m=0}^{1}(k-m)\left\{{n-m-1 \atop k-m}\right\}\Big) + \left\{{n-2 \atop k-2}\right\} = \cdots \\ &= \Big(\sum_{m=0}^{\ell-1}(k-m)\left\{{n-m-1 \atop k-m}\right\}\Big) + \left\{{n-\ell \atop k-\ell}\right\}.\end{aligned}$$

A more formal induction to prove this is also possible.

Exercise. Unfold (16) the other way.

Note: $b(n) = \sum_k \left\{{n \atop k}\right\}$ = total number of ways of partitioning X_n, the total number of equivalence relations on X_n.

$b(0) = b(1) = 1$, $b(2) = 2$, $b(3) = 5$, $b(4) = 15$, etc. There is no known simple formula for $b(n)$.

Stirling numbers of the first kind

Let g be a permutation of a set X. An *orbit* of g is a set of the form $\{g^k x \mid k \in \mathbb{Z}\}$ where $x \in X$. A *cycle* of g is a permutation of the form $g|_Y$ (restriction) where Y is an orbit of g.

For example, the orbits of $(123)(67) \in S_7$ are $\{1,2,3\}$, $\{4\}$, $\{5\}$, $\{6,7\}$. Its cycles are (123), (4), (5), (67). Cycles are always non-empty. Note: $(1234) = (2341)$ etc (cyclic permutation).

Definition 5: $\begin{bmatrix} n \\ k \end{bmatrix}$ is the number of permutations in S_n with k cycles.

Example 23: $n = 4$, $k = 2$. $(1,2,3)(4)$, $(1,3,2)(4)$, $(1,2,4)(3)$, $(1,4,2)(3)$, $(1,3,4)(2)$, $(1,4,3)(2)$, $(2,3,4)(1)$, $(2,4,3)(1)$, $(1,2)(3,4)$, $(1,3)(2,4)$, $(1,4)(2,3)$. So $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 11$

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \end{bmatrix} &= 1\,, \\ \begin{bmatrix} n \\ 0 \end{bmatrix} &= 0\,, \text{ for } n > 0 \\ \begin{bmatrix} n \\ 1 \end{bmatrix} &= (n-1)! \end{aligned}$$

the last equality because we can write the same n-cycle in precisely n different ways: given an n-cycle, choose as first digit any one of the n digits in the cycle, and then the cyclic order determines the order of the remaining digits.

Note that

$$\sum_k \begin{bmatrix} n \\ k \end{bmatrix} = n!$$

the total number of permutations of X_n.

Proposition 11 *Basic Recurrence Relation:*

$$\begin{bmatrix} n \\ k \end{bmatrix} = (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} \tag{17}$$

Proof
Either (n) is a cycle by itself, leaving $X_n \setminus \{n\} = X_{n-1}$ to be decomposed into $k-1$ disjoint cycles (in $\left[\begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix}\right]$ ways), or we decompose $X_n \setminus \{n\}$ into k non-empty cycles in $\left[\begin{smallmatrix} n-1 \\ k \end{smallmatrix}\right]$ ways and then insert the n into one of the cycles. For any given decomposition of X_{n-1} into disjoint cycles, there are $(n-1)$ ways of inserting n into one of them: put n immediately before any of the other $(n-1)$ numbers. □

n	$\left[{n \atop 0}\right]$	$\left[{n \atop 1}\right]$	$\left[{n \atop 2}\right]$	$\left[{n \atop 3}\right]$	$\left[{n \atop 4}\right]$	$\left[{n \atop 5}\right]$
0	1					
1	0	1				
2	0	1	1			
3	0	2	3	1		
4	0	6	11	6	1	
5	0	24	50	35	10	1

Example 24: How many ways are there of inserting 6 into $(1, 2, 3)(4, 5)$?

$$
\begin{array}{c}
(6, 1, 2, 3)(4, 5) \\
(1, 6, 2, 3)(4, 5) \\
(1, 2, 6, 3)(4, 5) \\
(1, 2, 3)(6, 4, 5) \\
(1, 2, 3)(4, 6, 5)
\end{array}
$$

(Note that $(1, 2, 3, 6)(4, 5)$ is the same as $(6, 1, 2, 3)(4, 5)$, and $(1, 2, 3)(4, 5, 6)$ is the same as $(1, 2, 3)(6, 4, 5)$, as decompositions into disjoint cycles — and, thus, of course, as permutations). Gives $6 - 1 = 5$ possible ways.

Exercise. $\displaystyle\sum_k (-1)^k \left[{n \atop k} \right] = 0$ if $n \geqslant 2$.

Recall the rising and falling powers from Assignment 2. We were able to express ordinary powers in terms both of rising powers, and of falling powers. For example

$$
\begin{aligned}
x^4 &= x^{\underline{4}} + 6x^{\underline{3}} + 7x^{\underline{2}} + x^{\underline{1}} \\
x^4 &= x^{\overline{4}} - 6x^{\overline{3}} + 7x^{\overline{2}} - x^{\overline{1}}
\end{aligned}
$$

Properties and examples about Stirling numbers

Theorem 12 *1.* $x^n = \sum_k \left\{ {n \atop k} \right\} x^{\underline{k}}$

2. $x^n = \sum_k (-1)^{n-k} \left\{ {n \atop k} \right\} x^{\overline{k}}$

3. $x^{\underline{n}} = \sum_k (-1)^{n-k} \left[{n \atop k} \right] x^k$

4. $x^{\overline{n}} = \sum_k \left[{n \atop k} \right] x^k$

Proof
We do (1) and (3). (2) and (4) are similar.

1. Induction on n. $n=0$ gives $1=1$, true. So assume true for $n-1$. Note that $x^{\underline{k+1}} = x^{\underline{k}}(x-k)$. So $x.x^{\underline{k}} = x^{\underline{k+1}} + kx^{\underline{k}}$. We have

$$\begin{aligned}
x^n &= x.x^{n-1} \\
&= x\left(\sum_k \left\{ \begin{array}{c} n-1 \\ k \end{array} \right\} x^{\underline{k}}\right) \text{ (by induction)} \\
&= \sum_k \left\{ \begin{array}{c} n-1 \\ k \end{array} \right\} (x^{\underline{k+1}} + kx^{\underline{k}}) \\
&= \sum_k \left\{ \begin{array}{c} n-1 \\ k \end{array} \right\} x^{\underline{k+1}} + \sum_k k \left\{ \begin{array}{c} n-1 \\ k \end{array} \right\} x^{\underline{k}} \\
&= \sum_k \left\{ \begin{array}{c} n-1 \\ k-1 \end{array} \right\} x^{\underline{k}} + \sum_k k \left\{ \begin{array}{c} n-1 \\ k \end{array} \right\} x^{\underline{k}} \\
&= \sum_k \left\{ \begin{array}{c} n \\ k \end{array} \right\} x^{\underline{k}} \text{(by Proposition 23)}
\end{aligned}$$

3. Induction on n. $n=0$ gives $1=1$, true. Assume true for $n-1$.

$$\begin{aligned}
x^{\underline{n}} &= x^{\underline{n-1}}(x-n+1) \\
&= (x-n+1)\sum_k \left[\begin{array}{c} n-1 \\ k \end{array} \right] (-1)^{n-1-k} x^k \text{ (by induction)} \\
&= \sum_k \left[\begin{array}{c} n-1 \\ k \end{array} \right] (-1)^{n-1-k} x^{k+1} + \sum_k \left[\begin{array}{c} n-1 \\ k \end{array} \right] (-1)^{n-k}(n-1)x^k \\
&= \sum_k \left[\begin{array}{c} n-1 \\ k-1 \end{array} \right] (-1)^{n-k} x^k + \sum_k \left[\begin{array}{c} n-1 \\ k \end{array} \right] (-1)^{n-k}(n-1)x^k \\
&= \sum_k \left[\begin{array}{c} n \\ k \end{array} \right] (-1)^{n-k} x^k \text{ (by Proposition 24)}
\end{aligned}$$

□

Example 25: Using parts (3) and (1) of theorem 12 yields

$$x^{\underline{n}} \overset{(3)}{=} \sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} x^k \overset{(1)}{=} \sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \Big(\sum_\ell \begin{Bmatrix} k \\ \ell \end{Bmatrix} x^{\underline{\ell}} \Big).$$

But the polynomials $\{x^{\underline{\ell}} : \ell \in \mathbb{Z}_{\geqslant 0}\}$ (falling powers) are linear independent. Therefore, we can compare coeffients of $x^{\underline{\ell}}$ which yields

$$[n=\ell] = \sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \begin{Bmatrix} k \\ \ell \end{Bmatrix}.$$

Example 26: Here is an example of an identity which can be obtained *bijectively*. Let $F(k)$ denote the set of permutations $g \in S_{n+1}$ of $\ell+1$ cycles such that $n+1$ is in a $(k+1)$-cycle, and $f(k)$ the number of elements of $F(k)$. So

$$\begin{bmatrix} n+1 \\ \ell+1 \end{bmatrix} = \sum_k f(k). \tag{18}$$

We choose an element of $F(k)$ as follows.

Step 1. Choose the orbit $\{g^m(n+1) \mid m \in \mathbb{Z}\}$ of $n+1$. The number of choices is $\binom{n}{k}$.

Step 2. Choose the cycle of $n+1$. The number of them is $\left[{k+1 \atop 1}\right] = k!$.

Step 3. Finish g. There are $\left[{n-k \atop \ell}\right]$ choices.

We are dealing with a homogeneous tree so

$$f(k) = \binom{n}{k} k! \begin{bmatrix} n-k \\ \ell \end{bmatrix}$$

so by (18) we find

$$\begin{bmatrix} n+1 \\ \ell+1 \end{bmatrix} = \sum_k \binom{n}{k} k! \begin{bmatrix} n-k \\ \ell \end{bmatrix}.$$

5.2 Harmonic Numbers

We define

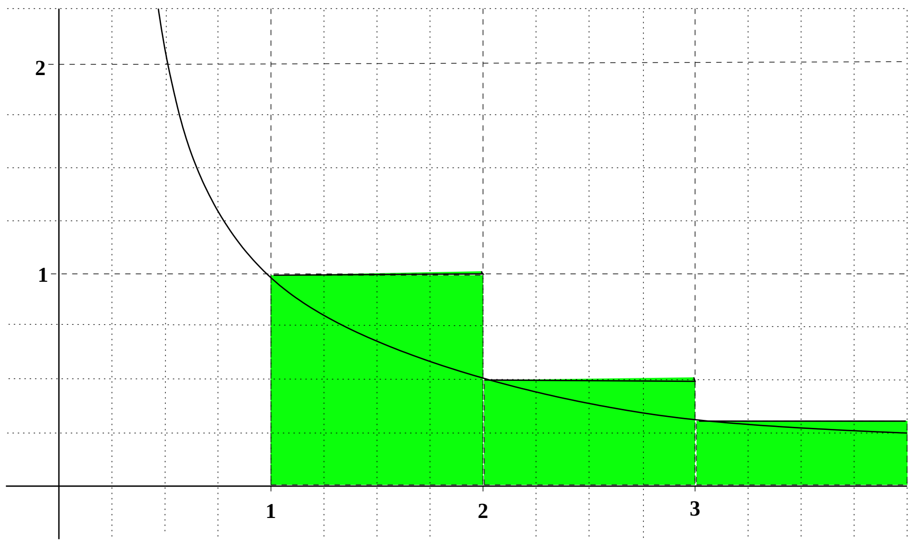
$$H_n = \sum_{k=1}^{n} \frac{1}{k}, \quad H_0 = 0.$$

Recall that

$$\log x = \int_1^x \frac{1}{x} dx$$

(where $\log x$ means $\log_e x$).

Consider the following figure, showing the graph of $y = 1/x$ and of a step function f whose integral between 1 and $n+1$ is H_n.

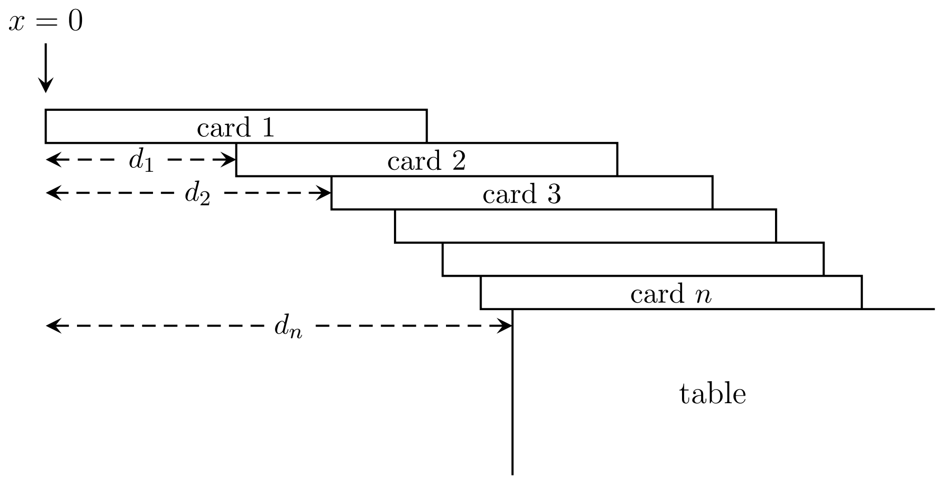


It is clear that $H_n > \log(n+1) > \log n$, and that $\log n > H_n - 1$ (throw away the first (square) box and shift all the others one unit to the left). Hence $H_n - \log(n+1)$ is bounded above by 1. It is an increasing sequence, and so tends to a limit, known as γ. Since $\log(n+1) - \log n = \log(1+1/n) \to 0$ as $n \to \infty$,

$$\lim_{n\to\infty}(H_n - \log n) = \gamma, \text{ which is approximately } \quad = 0.577215\ldots$$

Some problems in which H_n occur.

Example 27: Suppose we have a stack of n cards each of length 2 units and we try to balance them on edge of table to get maximal overhang. What is the largest possible overhang?



Let d_n be a possible overhang (possible in the sense that the cards don't fall). Number the cards 1–n from top to bottom. Since the first card does not fall, we have

$$d_1 \leqslant 1.$$

The reason for this is that the center of the first card must be above the second card.

More generally, the center of gravity of the first k cards must be above the $(k+1)$-st card, where the table plays the role of the $(n+1)$-st card. In order to compute this center of gravity, put $x = 0$ at the end of the first card. Then the center of gravity of the first k cards is

$$\frac{(1+d_0)+(1+d_1)+\cdots+(1+d_{k-1})}{k}$$

and therefore,

$$d_k \leqslant \frac{(1+d_0)+(1+d_1)+\cdots+(1+d_{k-1})}{k}. \tag{19}$$

Here $d_0 = 0$.

Note that, the greater $d_0, \ldots, d_k$ are, the greater the upper bound for d_k given by (19) is. It follows that d_n is greatest precisely when equality holds in (19), for all k:

$$d_k = \frac{(1+d_0)+(1+d_1)+\cdots+(1+d_{k-1})}{k} \qquad \text{for all } k,$$

or

$$kd_k = k + (d_0 + \cdots + d_{k-1}). \tag{20}$$

Replacing k by $k-1$ gives

$$(k-1)d_{k-1} = (k-1) + (d_0 + \cdots + d_{k-2}). \tag{21}$$

Substracting (20) and (21) gives

$$k\,d_k - (k-1)d_{k-1} = 1 + d_{k-1},$$

that is,

$$d_k = \frac{1}{k} + d_{k-1}.$$

Since $d_0 = 0$ we find that d_k is just the harmonic number $d_k = H_k$.

Since $H_n \to \infty$ as $n \to \infty$, we can get arbitrarily large overhang. $H_4 > 2$, so with 4 cards we can get the top card to be clear of the table. But note $H_{1000000} = 14.39$ so to get an overhang of 7 card lengths we need nearly 1000000 cards.

Example 28: You are collecting football stickers. The complete set has size n. You buy them one at a time, selected randomly from a complete set. How many do you expect to have to buy to get a full set?

Suppose you already have $k < n$ distinct stickers. How many do you expect to have to buy until you get a new sticker number $k+1$? At this point, the probability that a random sticker is new is $(n-k)/n = p$. The expected time for new sticker $k+1$ is $\sum_{r \geqslant 1} r\,P(r)$, where $P(r)$ is the probability that you get the new sticker on your r'th purchase. This occurs if you get $r-1$

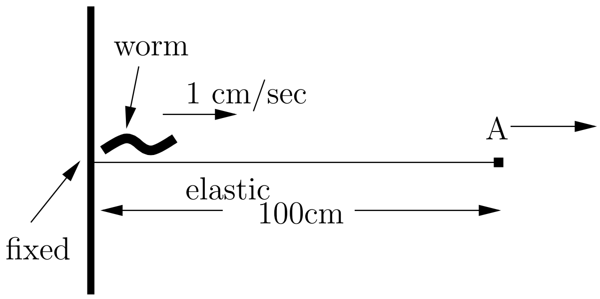
old stickers in a row, and then get a new one. So $P(r) = (1-p)^{r-1}p$, and expected time is

$$\sum_{r\geqslant 1} r\,(1-p)^{r-1}\,p = \sum_{r\geqslant 1} r\,q^{r-1}\,p = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} = \frac{n}{n-k}.$$

So total expected waiting time for full set is

$$\sum_{k=0}^{n-1} \frac{n}{n-k} = nH_n \sim n \ln n\,.$$

Example 29: Consider a worm on an elastic band. The worm starts at one end and crawls towards the other at 1cm per second. The elastic band initially has length 100cm, but after each second, it is stretched by 100cm.



Does worm ever reach end of elastic? After 1 second 1/100'th journey is over. Then the elastic is stretched to 200cm long. But worm remains 1/100'th way along since the stretching is uniform. In 2nd second, worm completes further 1/200'th of journey. So in total $1/100 + 1/200$ has been done. This remains true after stretching to 300cm. In 3rd second 1/300'th of journey done. After n seconds, the fraction of journey completed is $1/100+1/200+\ldots+1/100n = H_n/100$. Since $H_n \to \infty$ the answer to the question is yes.

It takes approximately n seconds where n is the first integer with $H_n \geqslant 100$. $H_n \sim \log n + \gamma$, so,

$$\begin{aligned} n \sim e^{100-\gamma} &= e^{99.423} \\ &= 4.79 \times 10^{35}\text{years} \end{aligned}$$

At this stage, length of elastic $\sim 10^{25}$ light years!

We can also do a continuous version: same but elastic is stretched continuously with far end moving at 100cm/s. Length of elastic is $l = 100t + 100$ where t is time. So we get

$$\frac{dx}{dt} = \frac{x}{t+1} + 1\,.$$

Solution $x = \log(t+1)(t+1)$.

$$\frac{x}{l} = \frac{\log(t+1)}{100}$$

$x = l$ when $\ln(t+1) = 100 \Rightarrow t = e^{100} - 1$.

5.3 Fibonacci Numbers, F_n

Definition 6: They are defined by $F_1 = F_2 = 1$, $F_n = F_{n-1} + F_{n-2}$. This defines F_n for all $n \in \mathbb{Z}$.

n	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
F_n	5	-3	2	-1	1	0	1	1	2	3	5	8	13

In general $F_{-n} = (-1)^{n-1}F_n$ (proof by (strong) induction on n).

Let
$$\phi = \frac{\sqrt{5}+1}{2} \sim 1.618$$
be a root of $x^2 - x - 1 = 0$. The other root is
$$\hat{\phi} = 1 - \phi = -\phi^{-1} \sim -0.618$$

Then
$$F_n = \frac{\phi^n - \hat{\phi}^n}{\sqrt{5}} = \left\lfloor \frac{\phi^n}{\sqrt{5}} + \frac{1}{2} \right\rfloor$$
which will be proved in chapter 6.

There are lots of identities involving F_n.

Cassini's Identity:
$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n$$
For example,
$$\begin{aligned} n &= 8\,, \qquad 13 \times 34 - 21^2 = 1 \\ n &= 6\,, \qquad 13 \times 5 - 64 = 1 \end{aligned}$$

Proof by induction.

Another identity:
$$F_{n+k} = F_kF_{n+1} + F_{k-1}F_n\,, \quad \text{for all } n, k \in \mathbb{Z}$$
by easy induction on k. For example,
$$\begin{aligned} k = n\,,\ F_{2n} &= F_n(F_{n+1} + F_{n-1})\,,\ \text{so } F_n \mid F_{2n} \\ k = 2n\,,\ F_{3n} &= F_{2n}F_{n+1} + F_{2n-1}F_n\,,\ \text{so } F_n \mid F_{3n} \end{aligned}$$

By easy induction $F_n \mid F_{kn}$ for all $k \geqslant 1$. In fact $\mathrm{hcf}(F_m, F_n) = F_{\mathrm{hcf}(m,n)}$ (harder).

Proposition 13 *Let $n \in \mathbb{N}$. Then n has a unique Fibonary representation*
$$n = F_{k_1} + F_{k_2} + \ldots + F_{k_r}\,,$$
with $k_i > k_{i+1} + 1$ and $k_r > 1$.

For example, $100 = 89 + 8 + 4 = F_{11} + F_6 + F_4$

Lemma 14 *If $r > 1$, then*

$$F_r + F_{r-2} + F_{r-4} + \ldots + F_{(3\ or\ 2)} < F_{r+1}$$

Proof
Induction on r

$$\begin{aligned} r &= 2: \ F_2 < F_3 \\ r &= 3: \ F_3 < F_4 \end{aligned}$$

For $r > 3$, by induction, $F_{r-2} + \ldots + F_{(3 \text{ or } 2)} < F_{r-1}$. So $F_r + (F_{r-2} + \ldots + F_{(3 \text{ or } 2)}) < F_r + F_{r-1} = F_{r+1}$. This proves the lemma. □

Proof of proposition
Existence: induction on n. For $n = 1$, $1 = F_2$, so OK.

For $n > 1$, choose k_1, with $F_{k_1} \leqslant n < F_{k_1+1}$. Apply induction to $n - F_{k_1}$. Since $n - F_{k_1} < F_{k_1+1} - F_{k_1} = F_{k_1-1}$. So we get $k_2 < k_1 - 1$ when we write $n - F_{k_1} = F_{k_2} + F_{k_3} + \ldots + F_{k_n}$. Since $F_2 = F_1 = 1$, we never need to use F_1, so get $k_r > 1$.

Uniqueness: to prove uniqueness let $n = F_{k_1} + \ldots + F_{k_r}$, $k_i > k_{i+1} + 1$, $k_r > 1$. k_1 must be the largest possible such that

$$F_{k_1} \leqslant n < F_{k_1+1}\,.$$

Otherwise, by lemma, sum on right is less than n. Then apply induction to $n - F_{k_1}$. The $k_2, \ldots, k_r$ are uniquely determined. □

Example 30: What is the Fibonary expansion of $3F_n$ ($n \geqslant 0$)? We begin by computing a few small cases.

n	F_n	$3F_n$	Fibonary expansion of $3F_n$
0	0	0	0
1	1	3	F_4
2	1	3	F_4
3	2	6	$F_5 + F_2$
4	3	9	$F_6 + F_2$
5	5	15	$F_7 + F_3$

From the last two lines we guess

$$3F_n = F_{n+2} - F_{n-2}. \tag{22}$$

This is true for all $n \in \mathbb{Z}$ and can easily be proved by induction. Now (22) is the Fibonary expansion for $3F_n$ provided $n \geqslant 4$. For $0 \leqslant n < 4$ refer to the table.

Example 31: Game: two players A and B. A chooses $n \in \mathbb{N}$, $n > 1$. Then players take it in turn to subtract an integer m (which they choose each time) from n. The player who first gets it to 0 wins. Rules:

1. B must not subtract n on first go.
2. You may never subtract more than $2m$, where m was previous number subtracted.

For example, when $n = 12$,

n	B	A	B	A	B	A	B
	-2	-3	-1	-1	-1	-1	-2
12	10	7	5	4	3	2	0

Here B wins.

Winning strategy: A should choose a Fibonacci Number. Otherwise B should write

$$n = F_{k_1} + \ldots + F_{k_r}$$

as in the proposition and subtract F_{k_r}.

For example, $n = 12$, $12 = 8 + 3 + 1$. B subtracts 1 to get 11. In the game above A could have won. $10 = 8 + 2$. A subtracts 2 to get Fibonacci number 8.

6 Generating Functions

6.1 Basic Manipulation

Let $G = \{g_0, g_1, g_2, \ldots\}$ be an infinite sequence, $g_i \in \mathbb{C}$. Let $g_n = 0$ for $n < 0$. The generating function of the sequence (g_i) is defined to be

$$G(z) = \sum_{n \in \mathbb{Z}} g_n z^n = g_0 + g_1 z + g_2 z^2 + g_3 z^3 \ldots .$$

We will nearly always treat this as a formal power series, and not concern ourselves with convergence.

Basic Operations:

1. *Linear sums*: Let $F = (f_n)$ and $G = (g_n)$ be sequences; $\alpha, \beta \in \mathbb{R}$. Then

$$(\alpha F + \beta G)(z) = \sum_n (\alpha f_n + \beta g_n) z^n .$$

2. *Shifting*: $(m \in \mathbb{N})$

$$z^m G(z) = \sum_n g_n z^{n+m} = \sum_n g_{n-m} z^n .$$

3. *Scalar Multiplication*:

$$G(cz) = \sum_n g_n c^n z^n \quad (\neq cG(z))$$

4. *Differentiation*:

$$\begin{aligned} G'(z) &= \sum_n (n+1) g_{n+1} z^n \\ zG'(z) &= \sum_n n g_n z^n . \end{aligned}$$

5. *Multiplication or convolution*:

$$\begin{aligned} F(z)G(z) &= \sum_n \left(\sum_k f_k g_{n-k} \right) z^n \\ &= f_0 g_0 + (f_0 g_1 + f_1 g_0) z + (f_0 g_2 + f_1 g_1 + f_2 g_0) z^2 + \ldots \end{aligned}$$

6. *Division*: Since

$$(1-z)(1 + z + z^2 + z^3 + \ldots) = 1 ,$$

we can write

$$\frac{1}{1-z} = \sum_{n \geqslant 0} z^n .$$

In general

$$\frac{G(z)}{1-z} = \left(\sum_n g_n z^n\right)(1+z+z^2+\ldots) = \sum_n \left(\sum_{k\leqslant n} g_k\right) z^n$$

Dividing by $(1-z)$ replaces a sequence by the sequence of its partial sums. For example, $G = (1,1,1,1,\ldots)$,

$$\begin{aligned} G(z) &= \frac{1}{1-z} \\ \frac{1}{(1-z)^2} &= \sum_n (n+1) z^n \end{aligned}$$

is the generating function of $(1,2,3,4,\ldots)$ (also get by differentiating $1/(1-z)$)

7. The generating function of $(1, \frac{1}{1!}, \frac{1}{2!}, \frac{1}{3!}, \ldots)$ is

$$e^z = \sum_{n\geqslant 0} \frac{z^n}{n!}.$$

The generating function of $(0, \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots)$ is

$$\ln\left(\frac{1}{1-z}\right) = \sum_{n\geqslant 1} \frac{z^n}{n}.$$

8. Strictly speaking, the above should be taken as *definitions.* They are different from the things (definitions or theorems) you've seen in analysis, however alike they seem. All results you know in other settings are true and should, strictly speaking be proved, but we shan't. To give an idea what sort of properties we mean — there are many — here are a few:

$$\begin{aligned} z^m(z^n G(z)) &= z^{m+n} G(z) && \\ (FG)H &= F(GH) && \text{for all GF's } F, G, H, \\ (FG)' &= F'G + FG' && \text{for all GF's } F, G, \qquad (23) \\ \frac{d}{dz} e^z &= e^z. && \end{aligned}$$

Exercise. Prove (23).

6.2 Representations of sequences

There are many ways to define a sequence $a_0, a_1, \ldots$ of complex numbers (usually natural numbers for us). Here are the ones we're interested in, illustrated by the example where $a_n = 2^n$.

(1) **Cardinality.** Give a set B_n such that $a_n = |B_n|$. In our example, $B_n = \{1,2\}^n$.

(2) **Recursion.** A formula which expresses each a_n in the previous terms. For example $a_n = 2a_{n-1}$ or

$$a_n = a_{n-1} + a_{n-2} + \cdots + a_0 + 1.$$

(3) **Equation for the GF.** An equation involving the generating function $A(x) = \sum_{n\geqslant 0} a_n x^n$. We distinguish between the following.

(3a) **Algebraic equation.** For example

$$(1-2x)^2 A(x)^2 + (1-2x)A(x) = 2.$$

(3b) **Differential equation.** This is an equation involving at least one of $A'(x)$, $A''(x)$, ... (derivatives) and possibly $A(x)$. For example

$$A'(x) = \frac{2\,A(x)}{1-2x}.$$

(3c) **Functional equation.** Different arguments of the GF are combined in one equation. For example

$$2\,A(2x^2) = A(x) + A(-x).$$

(4) **Formula for GF.** For example

$$A(x) = \frac{1}{1-2x}.$$

(5) **Partial fraction expansion for GF.** For example (again)

$$A(x) = \frac{1}{1-2x}.$$

(6) **Formula for a_n.** In our example, $a_n = 2^n$.

Many problems are of the following sort: a sequence a_n is given by one of the above methods. You're then asked to translate it into a specified other method. Our favourite line is

$$(1) \to (2) \to (3) \to (4) \to (5) \to (6)$$

but anything else is also possible.

6.3 Partial Fractions

You're supposed to know about partial fractions, but for all clarity we give the main result and a few examples.

In this section we are interested in the steps

(4) Formula for GF $\to$ (5) Partial fraction $\to$ (6) Formula for a_n

By (4) we are given a direct formula for a GF. We suppose it's a rational function (a quotient of two polynomials) because only then (5) makes sense.

By $\mathbb{C}[z]$ we denote the set of *polynomials* $a_0 + a_1 z + \cdots + a_n z^n$ in one variable z. The *degree* of this polynomial is n, if $a_n \neq 0$. The degree of the zero polynomial is -1. In $\mathbb{C}[z]$, we can add, substract and multiply; but the quotient of two polynomials is not always again a polynomial.

By $\mathbb{C}(z)$ we shall denote the set of *rational functions*, that is, quotients of polynomials.

In $\mathbb{C}(z)$ one can add, substract, multiply and divide by nonzero elements. In particular, it is a vector space over $\mathbb{C}$. The following proposition gives a basis.

Proposition 15 *Partial Fractions:*

(a) $\mathbb{C}(z)$ *has the basis*

$$\left\{z^k \;\middle|\; 0 \leqslant k\right\} \cup \left\{(z-\alpha)^{-k} \;\middle|\; \alpha \in \mathbb{C},\ 0 < k\right\}.$$

(b) Let $q \in \mathbb{C}[z]$ *be a nonzero polynomial, and let* $d \geqslant 0$*. Let* $V = V(q,d) \subset \mathbb{C}(z)$ *be the set of rational functions whose denominator is* q *(or a divisor of it), and such that the degree of the numerator is smaller than* d *plus the degree of the denominator. Here is a basis for the complex vector space* $V(q,d)$*, which is indeed a subset of the basis of (a):*

$$\left\{z^k \;\middle|\; 0 \leqslant k < d\right\} \cup \left\{(z-\alpha)^{-k} \;\middle|\; 0 < k,\ (z-\alpha)^{-k} q \in \mathbb{C}[z]\right\}.$$

Expressing a given rational function in its partial fraction simply means writing it as a linear combination of the particular basis given in part (a) of the above proposition. In practice, one calculates a partial fractions expansion by choosing q and d such that the involved rational function is in $V(q,d)$ – usually one immediately knows such q and d. Then one needs only consider the finite basis of $V(q,d)$ given in part (b).

Example 32: Find the partial fraction expansion of

$$G(x) = \frac{-25x}{(1-2x)^2(1+3x)}.$$

Solution. By Proposition 15 the answer must be of the form

$$\frac{-25x}{(1-2x)^2(1+3x)} = \frac{a}{1-2x} + \frac{b}{(1-2x)^2} + \frac{c}{1+3x}$$

where a, b, c are constants yet to be found. On multiplying both sides with the denominator $(1-2x)^2(1+3x)$ we find

$$\begin{aligned} -25x &= a(1-2x)(1+3x) + b(1+3x) + c(1-2x)^2 \\ &= a(1+x-6x^2) + b(1+3x) + c(1-4x+4x^2) \\ &= (a+b+c) + (a+3b-4c)x + (-6a+4c)x^2 \end{aligned}$$

so

$$\begin{cases} a+b+c=0 \\ a+3b-4c=-25 \\ -6a+4c=0. \end{cases}$$

You know how to solve this and you find $(a,b,c)=(2,-5,3)$, so

$$G(x)=\frac{2}{1-2x}+\frac{-5}{(1-2x)^2}+\frac{3}{1+3x}. \tag{24}$$

The second and last step is to turn partial fraction expansions into a direct formula for the coefficients. This is aided by the following formula.

Proposition 16 *Let $n \geqslant 0$. Then*

$$\frac{1}{(1-x)^n}=\sum_{k\geqslant 0}\binom{n+k-1}{k}x^k.$$

Proof. We have

$$(1-x)^{-n}=\sum_{k\geqslant 0}\binom{-n}{k}(-x)^k=\sum_{k\geqslant 0}\binom{n+k-1}{k}x^k.$$

The first identity is the Binomial Theorem 6, the second identity is upper negation (14). □

Example 33: Compute the coefficients in the GF (24).

Solution. By Proposition 16 we have

$$\begin{aligned} G(x)&=\frac{2}{1-2x}+\frac{-5}{(1-2x)^2}+\frac{3}{1+3x} \\ &=\sum_{k\geqslant 0}\Big(2(2x)^k-5(k+1)(2x)^k+3(-3x)^k\Big) \end{aligned}$$

so the coefficient of x^k is

$$2^{k+1}-5(k+1)2^k-3^{k+1}.$$

Remark. This remark does not belong to the syllabus, but you may find it helpful because it explains what power series of rational functions look like, and in practice many generating functions are rational functions. Let $G(z)$ be the generating function of $(g_0, g_1, \ldots)$. Then the following can be shown to be equivalent:

1. $G(z)$ is a rational function.

2. There is a recurrence $a_0 g_n + a_1 g_{n-1} + \ldots + a_k g_{n-k} = 0$ ($a_i \in \mathbb{C}$, $a_0 \neq 0$), which is true for all n big enough.

3. There are nonzero $c_i \in \mathbb{C}$ (i in some finite set I) and polynomials $p_i(x) \in \mathbb{C}[x]$ such that for n big enough, g_n is given by

$$g_n = \sum_{i \in I} p_i(n)\, c_i^n.$$

Exercise: Prove this, using Proposition 15 on partial fractions.

Exercise. Use partial fractions to evaluate the sum $\displaystyle\sum_{k=1}^{n} \frac{(-1)^k k}{4k^2 - 1}$.

6.4 Solving Recurrences

The examples in this subsection have in common that the recursion formula for the sequence $(g_n)_n$ is *linear inhomogeneous*, that is, of the form

$$g_n = \text{linear combination of the previous ones} + c_n$$

where c_n is known (and simple).

The procedure for finding a formula for g_n from the recursion formula is as follows.

Step 1. Write down a single equation for a recurrence, true for all n.

Step 2. Multiply by z^n and sum over n. This gives an equation in $G(z)$.

Step 3. Solve the equation to get an expression for $G(z)$.

Step 4. Solution g_n is coefficient of z^n in $G(z)$.

Example 34: Fibonacci Numbers:

$$g_0 = 0\,,\ g_1 = 1\,,\ g_n = g_{n-1} + g_{n-2}$$

But now we must let $g_n = 0$ for all $n < 0$.

Step 1:

$$g_n = g_{n-1} + g_{n-2}\,,\ \text{for } n \geqslant 2 \tag{25}$$

What happens for $n < 2$?

$$\begin{aligned} n = 1, &\qquad 1 = g_1 = g_0 + g_{-1} + 1 \\ n = 0, &\qquad 0 = g_0 = g_{-1} + g_{-2} \end{aligned}$$

In fact, (25) is true for all $n \leqslant 0$. Wrong only for $n = 1$. Single equation is

$$g_n = g_{n-1} + g_{n-2} + [n = 1]\,, \qquad \text{for all } n \in \mathbb{Z}\,.$$

Step 2:

$$g_n z^n = g_{n-1} z^n + g_{n-2} z^n + [n=1] z^n$$

Sum over $n \in \mathbb{Z}$ to get

$$G(z) = zG(z) + z^2 G(z) + z\,.$$

Step 3:

$$G(z) = \frac{z}{1 - z - z^2}$$

Now factorise denominator and use partial fractions. Note: if $x^2 + ax + b = 0$ has roots α, β then $1 + az + bz^2 = (1 - \alpha z)(1 - \beta z)$. In this case, the roots of $x^2 - x - 1 = 0$ are ϕ and $\hat{\phi}$ where $\phi = (1 + \sqrt{5})/2$. So $1 - z - z^2 = (1 - \phi z)(1 - \hat{\phi} z)$. By partial fractions,

$$\frac{z}{(1 - \phi z)(1 - \hat{\phi} z)} = \frac{1}{\sqrt{5}} \left(\frac{1}{1 - \phi z} - \frac{1}{1 - \hat{\phi} z} \right)$$

since $\sqrt{5} = \phi - \hat{\phi}$.

Step 4: Coefficient of z^n in $1/(1 - \phi z)$ is ϕ^n for all $n \geqslant 0$. So $g_n = (\phi^n - \hat{\phi}^n)/\sqrt{5}$

Example 35: $g_0 = g_1 = 1$

$$g_n = g_{n-1} + 2g_{n-2} + (-1)^n \quad (n \geqslant 2)$$

n	0	1	2	3	4	5	6	7
g_n	1	1	4	5	14	23	52	97

Step 1:

$$\begin{aligned} n &= 1\,, & 1 &= g_1 = g_0 + 2g_{-1} + (-1)^n + 1 \\ n &= 0\,, & 1 &= g_0 = g_{-1} + 2g_{-2} + (-1)^n \\ n &< 0\,, & g_n &= g_{n-1} + 2g_{n-2} \end{aligned}$$

So single equation:

$$g_n = g_{n-1} + 2g_{n-2} + (-1)^n [n \geqslant 0] + [n = 1] \quad \text{for all } n \in \mathbb{Z}$$

Step 2:

$$G(z) = zG(z) + 2z^2 G(z) + \sum_{n \geqslant 0} (-1)^n z^n + z$$

Step 3: Note that
$$\sum_{n\geqslant 0}(-1)^n z^n = \frac{1}{1+z}\,.$$

$$\begin{aligned} G(z) &= \frac{1+z(1+z)}{(1+z)(1-z-2z^2)} \\ &= \frac{1+z+z^2}{(1-2z)(1+z)^2} \\ &= \frac{A}{1-2z} + \frac{B}{(1+z)} + \frac{C}{(1+z)^2} \end{aligned}$$

Solve for A, B, C by comparing coefficients to get $A = 7/9\,,\ B = -1/9\,,\ C = 1/3$.
$$G(z) = \frac{7}{9(1-2z)} - \frac{1}{9(1+z)} + \frac{1}{3(1+z)^2}\,.$$

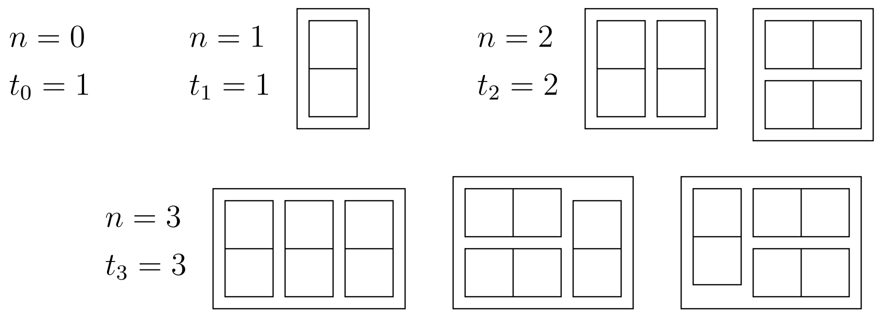
Step 4: Expressions for each of the three terms can easily be obtained from the Binomial Theorem. Adding them, we find: coefficient of z^n is

$$\frac{7}{9}2^n - \frac{1}{9}(-1)^n + \frac{n+1}{3}(-1)^n = \frac{7\cdot 2^n}{9} + \left(\frac{n}{3}+\frac{2}{9}\right)(-1)^n$$

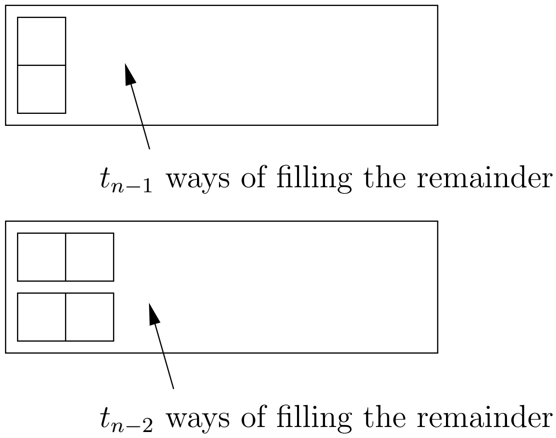
For example $n = 4$, $7 \times 16/9 + 14/9 = 14$

Example 36:

Part 1. How many ways can you tile a $2 \times n$ rectangle with n dominoes (1×2 rectangles). Call this t_n.

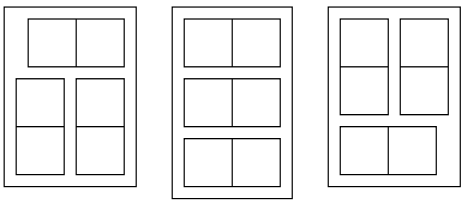


For recurrence pattern at left is either



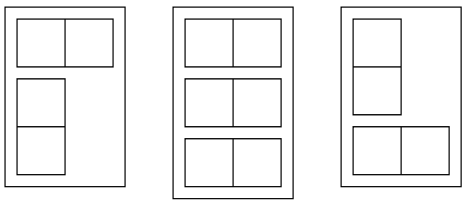
$$t_n = t_{n-1} + t_{n-2} \Rightarrow t_n = F_{n+1}$$

Part 2. Let u_n be the number of ways of tiling a $3 \times n$ rectangle with 1×2 dominoes. $u_0 = 1$, $u_1 = 0$, in fact $u_n = 0$ for n odd since $3n$ is odd.

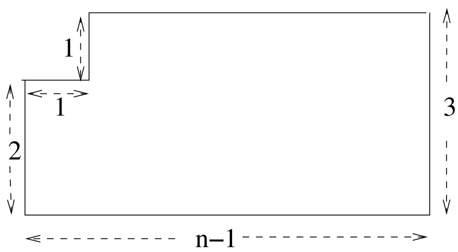


$u_2 = 3$

Pattern on left starts in one of the following 3 ways:

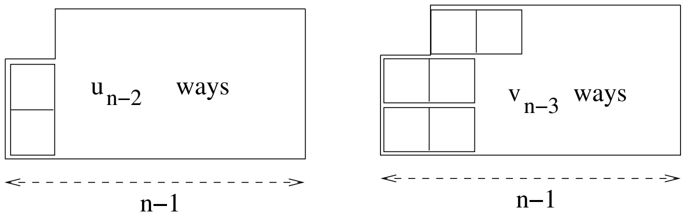


Let v_{n-1} be the number of ways of completing a tiling of the $3 \times n$ rectangle after starting on the left in the first (or, symmetrically, the third) of these ways. That is, v_{n-1} is the number of ways of tiling the shape



Then $u_n = u_{n-2} + 2v_{n-1}$ (with $u_0 = 1\,,\ u_1 = 1$).

Moreover, as the following figure shows, $v_{n-1} = u_{n-2} + v_{n-3}$,



i.e. $v_n = u_{n-1} + v_{n-2}$, for $n \geqslant 2$.

Evidently $v_0 = 0$, $v_1 = 1$.

Let u, v be generating functions for (u_n), (v_n).

Step 1: Equations valid for all n:

$$\begin{aligned} u_n &= u_{n-2} + 2v_{n-1} + [n = 0] \\ v_n &= u_{n-1} + v_{n-2} \end{aligned}$$

Step 2:

$$U(z) = z^2 U(z) + 2zV(z) + 1 \tag{26}$$

$$V(z) = zU(z) + z^2 V(z) \tag{27}$$

Step 3: (27) gives

$$V(z) = \frac{zU(z)}{1 - z^2}.$$

Substitute into (26):

$$\begin{aligned} U(z) &= \frac{2z^2 U(z)}{1 - z^2} + z^2 U(z) + 1 \\ U(z) &= \frac{1 - z^2}{1 - 4z^2 + z^4} = \frac{1 - w}{1 - 4w + w^2}, \text{ where } w = z^2 \end{aligned}$$

Roots of $x^2 - 4x + 1$ are $2 \pm \sqrt{3}$. So

$$\begin{aligned} U(z) &= \frac{1 - w}{(1 - (2 + \sqrt{3})w)(1 - (2 - \sqrt{3})w)} \\ &= \frac{3 + \sqrt{3}}{6} \cdot \frac{1}{1 - (2 + \sqrt{3})w} + \frac{3 - \sqrt{3}}{6} \cdot \frac{1}{1 - (2 - \sqrt{3})w} \end{aligned}$$

(by partial fractions). Note that

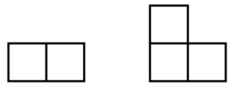
$$\frac{3 + \sqrt{3}}{6} = \frac{1}{3 - \sqrt{3}}.$$

Step 4: Coefficient of z^n

$$u_n = \frac{(2 + \sqrt{3})^n}{3 - \sqrt{3}} + \frac{(2 - \sqrt{3})^n}{3 + \sqrt{3}} \quad (\text{so } u_{2n+1} = 0)$$

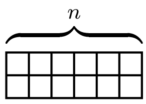
n	0	1	2	3	4	5	6	7	8
u_n	1	0	3	0	11	0	41	0	153
v_n	0	1	0	4	0	15	0	56	0

Example 37: Let p_n denote the number of ways to tile a $2 \times n$ rectangle by dominoes and hooks. Find a recurrence formula for p_n.



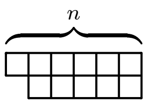
domino hook

Solution. We can write

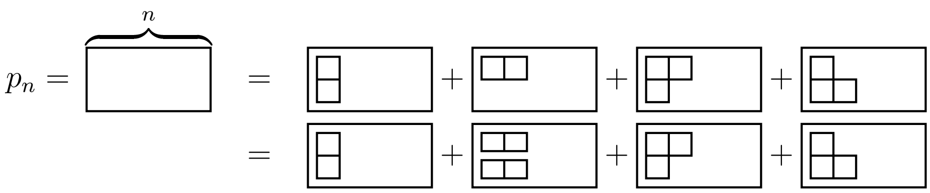
$p_n =$ 

if we agree that a picture stands for a number, namely the number of ways to tile the picture by dominoes and hooks. Or the number of ways to finish the tiling if a tiling has already begun (this applies in the pictures starting from (28)).

We also define

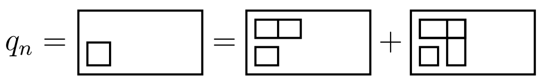
$q_n =$ .

By considering all possibilities to cover the top left box, we find



$$= p_{n-1} + p_{n-2} + 2q_{n-1}. \tag{28}$$

Also



$$= q_{n-1} + p_{n-2}. \tag{29}$$

Equation (28) helps us express q_n in terms of the p_i:

$$2q_{n-1} = p_n - p_{n-1} - p_{n-2}.$$

Inserting this into (29) we find

$$\begin{aligned} 0 &= 2q_n - 2q_{n-1} - 2p_{n-2} \\ &= (p_{n+1} - p_n - p_{n-1}) - (p_n - p_{n-1} - p_{n-2}) - 2p_{n-2} \\ &= p_{n+1} - 2p_n - p_{n-2}. \end{aligned}$$

Example 38: Given a product $x_0 x_1 x_2 \dots x_n$ of $n+1$ terms. How many different ways can we bracket this? i.e. how many different ways can we calculate it using n multiplications? Call it c_n.

$$c_0 = c_1 = 1$$

$$\begin{aligned}
n = 2 \qquad & (x_0 x_1) x_2 \text{ or } x_0(x_1 x_2) \\
\Rightarrow \quad & c_2 = 2 \\
n = 3 \qquad & x_0(x_1(x_2 x_3)),\ x_0((x_1 x_2)x_3),\ (x_0 x_1)(x_2 x_3) \\
& \quad ((x_0 x_1)x_2))x_3,\ x_0(x_1 x_2)x_3 \\
\Rightarrow \quad & c_3 = 5
\end{aligned}$$

To get a recurrence, suppose the final multiplication occurs between x_k and x_{k+1}

$$\Big((x_0 \dots)(\dots x_k)\Big) \times \Big((x_{k+1} \dots)(\dots x_n)\Big)$$

There are c_k ways of bracketing the first term and c_{n-k-1} ways of bracketing the second term. So

$$c_n = \sum_{k=0}^{n-1} c_k c_{n-k-1} = c_0 c_{n-1} + c_1 c_{n-2} + \dots + c_{n-1} c_0.$$

Step 1: For $n = 0$, $c_0 = 1$, $RHS = 0$, so

$$c_n = \sum_k c_k c_{n-1-k} + [n = 0]\,, \text{ for all } n \in \mathbb{Z}\,.$$

Step 2: Let $C(z)$ be the generating function of (c_n). Then

$$\begin{aligned}
C(z) &= \sum_n \left(\sum_k c_k c_{n-1-k} \right) z^n + 1 = z \sum_n \left(\sum_k c_k c_{n-1-k} \right) z^{n-1} + 1 \\
&= z \sum_n \left(\sum_k c_k c_{n-k} \right) z^n + 1 = z C(z)^2 + 1
\end{aligned}$$

(we have used convolution of sequences, (5) in the Basic Operations, to get the last line here).

Hence

$$z C(z)^2 - C(z) + 1 = 0,$$

and

Step 3:

$$C(z) = \frac{1 \pm \sqrt{1 - 4z}}{2z}$$

With positive sign get term $1/z$ in expansion, so that cannot be correct. So negative sign must be correct.

$$\begin{aligned}\sqrt{1-4z} &= (1-4z)^{1/2} \\ &= \sum_{k\geqslant 0}\binom{1/2}{k}(-4z)^k\,,\ \text{by binomial theorem} \\ &= 1+\sum_{k\geqslant 1}\frac{1}{2k}\binom{-1/2}{k-1}(-4z)^k\,,\ \text{by absorption.}\end{aligned}$$

So

$$\begin{aligned}C(z) &= \frac{1-\sqrt{1-4z}}{2z} \\ &= \sum_{k\geqslant 1}\frac{1}{k}\binom{-1/2}{k-1}(-4z)^{k-1} = \sum_{k\geqslant 0}\binom{-1/2}{k}\frac{(-4z)^k}{k+1}.\end{aligned}$$

Note that

$$\begin{aligned}\binom{-1/2}{k} &= \frac{-1/2\times -3/2\times -5/2\times\ldots\times(-2k-1)/2}{k!} \\ &= \frac{(-1)^k}{2^k k!}(1\times 3\times 5\times\ldots\times(2k-1)) \\ &= \frac{(-1)^k}{2^k k!}\frac{(2k)!}{2\times 4\times 6\ldots\times 2k} \\ &= \frac{(-1)^k(2k)!}{2^k k! 2^k k!} = \frac{(-1)^k}{4^k}\binom{2k}{k} = \frac{1}{(-4)^k}\binom{2k}{k}.\end{aligned}$$

So

$$C(z) = \sum_{k\geqslant 0}\binom{2k}{k}\frac{z^k}{k+1}.$$

Step 4:

$$c_n = \binom{2n}{n}\frac{1}{n+1}$$

These are called the Catalan Numbers.

n	0	1	2	3	4	5	6	7	8	9
c_n	1	1	2	5	14	42	132	429	1430	4862

$$\lim_{n\to\infty}\frac{c_{n+1}}{c_n} = 4$$

Another example of where these numbers arise: How many solutions are there of:

$$\begin{aligned}a_1+a_2+\ldots+a_{2n} &= 0, \quad a_i = \pm 1 \\ a_1+a_2+\ldots+a_j &\geqslant 0, \quad 0\leqslant j\leqslant 2n\end{aligned}$$

Call this number d_n. Let $S_{2k} = a_1 + a_2 + \ldots + a_{2k}$. Let k be maximal with $k < n$ and $S_{2k} = 0$ (possibly $k = 0$). There are d_k possibilities for $a_1, \ldots, a_{2k}$, we must have $a_{2k+1} = 1$ and $a_{2n} = -1$.

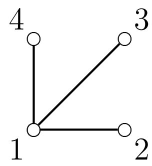
d_{n-k-1} possibilities for $a_{2k+2}, \ldots, a_{2n-1}$. So we get recurrence $d_n = \sum_k d_k d_{n-k-1}$ ($n \geqslant 1$). Since this is the same recurrence as for (c_n), and the initial values $c_1 = d_1$, $c_2 = d_2$ are the same, $c_n = d_n$ for all n.

Example 39: Spanning trees.

To avoid all confusion, let us recall a few definitions about graphs. A *graph* is a pair (V, E) where V is a set and E a subset of

$$V^{(2)} = \{\{x, y\} \mid x, y \in V,\ x \neq y\}.$$

The elements of V are the vertices, the elements of E the edges. In a picture, the elements of V are depicted by points and an edge $\{x, y\}$ by a line connecting the vertices x and y. Thus, the graph $(\{1, 2, 3, 4\}, \{\{1, 2\}, \{1, 3\}, \{1, 4\}\})$ can be depicted as follows.

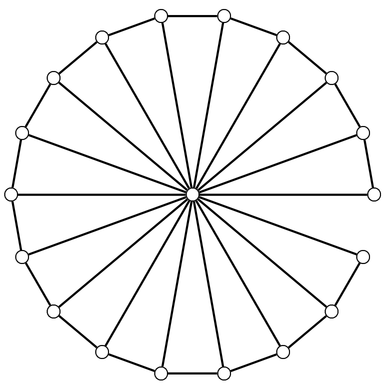


A graph (E, V) is *connected* if any two of its vertices x, y can be connected by a *path*

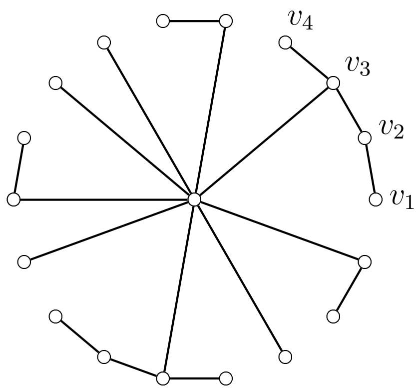
$$x = z_0, z_1, \ldots, z_k = y$$

which means that $\{z_i, z_{i+1}\}$ is an edge for all i. This path is called a *cycle* if $z_0, \ldots, z_{k-1}$ are all different but $z_k = z_0$, and $k > 2$.

A *tree* is a connected graph without cycles. A *spanning tree* of a graph (V, E) is a tree (V, T) with $T \subset E$.



The almost-wheel G_{18}



One of its spanning trees

For integers $n \geqslant 0$, let G_n denote the *almost wheel* shown in the picture. It has $n + 1$ vertices, one of which is depicted in the middle. Note that one of the 'edges' on the boundary is missing. On the right hand side, we see one of its spanning trees.

Let v_1, ..., v_n denote the vertices of G_n different from the origin, in this order starting and ending at the missing edge.

Let t_n denote the number of spanning trees of G_n. In order to find a formula for t_n, we begin with a recurrence formula.

The spanning tree in the picture is said to be of type 4; more generally, we call a spanning tree of G_n of type k if $\{v_i, v_{i+1}\}$ is an edge for $1 \leqslant i < k$ but not for $i = k$. The result is that $v_1, \ldots, v_k$ are connected along the boundary with each other but not with v_{k+1}.

In order to count the number of type k spanning trees, note that they depend on two things. Firstly, one among $v_1, \ldots, v_k$ needs to be connected to the origin: there are k choices for this. Secondly, all vertices except v_1, ..., v_k form a smaller almost-wheel G_{n-k}. A spanning tree for the G_{n-k} should be chosen for which there are t_{n-k} choices.

So there are $k\, t_{n-k}$ spanning trees of G_n of type k, and

$$t_n = \sum_{k=1}^{n} k\, t_{n-k} = \sum_{k=0}^{n} k\, t_{n-k} \quad \text{if } n \geqslant 1.$$

Let us write

$$T(z) = \sum_{n \geqslant 0} t_n z^n = \text{generating function of } t_n.$$

Then

$$\begin{aligned} T(z) &= 1 + \sum_{n \geqslant 1} t_n\, z^n = 1 + \sum_{n \geqslant 0} z^n \sum_{k=0}^{n} k\, t_{n-k} \\ &= 1 + \Big(\sum_{k \geqslant 0} k\, z^k\Big)\Big(\sum_{\ell \geqslant 0} t_\ell\, z^\ell\Big) = 1 + \frac{z}{(1-z)^2} T(z). \end{aligned}$$

It is easy to solve this for $T(z)$:

$$T(z) = \frac{1}{1 - \frac{z}{(1-z)^2}} = \frac{(1-z)^2}{1 - 3z + z^2}.$$

Unfortunately it is beyond our scope to look at the so-called *matrix-tree theorem* which gives the number of spanning trees of any graph as the determinant of a matrix.

6.5 The simplest sort of differential equations

Let f be a function of one variable, which may be a function of a real number, a function of a complex number, or a generating function. An *ordinary differential equation* is an equation involving one of f', f'', ... (derivatives) and possibly f. Usually the function f in a differential equation is unknown, the aim being to say something about the functions solving it.

We will be interested in differential equations of the form

$$f'(x) = a(x)\, f(x) \tag{30}$$

where $a(x)$ is an (explicitly) given function, and one looks for solutions f. Such equations are called linear homogeneous ordinary differential equations of order 1.

Proposition 17 *Let $A(x)$ denote a primitive of $a(x)$ (that is, $a(x) = A'(x)$). Then all solutions of (30) are given by*

$$f(x) = c\, e^{A(x)} \tag{31}$$

for a constant c.

Warning. This proposition needs to be taken with a grain of salt. For example, some generating functions are not the exponential of any generating function.

Proof. It is straightforward to show that the functions (31) are solutions to (30). We will now show that the converse also holds. It is also a way of reproducing the proposition, if you forget.

Suppose $f(x)$ satisfies (30). First we divide both sides by $f(x)$:

$$\frac{f'(x)}{f(x)} = a(x).$$

Then we note that a primitive of the left hand side is known:

$$\big(\log f(x)\big)' = a(x).$$

Now we use that $A(x)$ is a primitive of $a(x)$. Primitives are unique up to additive constants so we may write

$$\log f(x) = A(x) + b.$$

Taking the exponential of both sides gives (31) with $c = e^b$. □

A typical example of how to apply the proposition to generating functions is example 42 on Bell numbers.

6.6 Exponential Generating Functions

The EGF (Exponential Generating Function) for the sequence $(g_0, g_1, g_2, \ldots)$ is

$$\hat{G}(z) = \sum_{n \geqslant 0} \frac{g_n z^n}{n!}$$

For example $(1, 1, 1, \ldots)$ has EGF e^z, whereas it has GF $(1 - z)^{-1}$.

This is only a minor variation on the definition of ordinary Generating Function, but in some situations it is a lot more convenient. Some recurrences are much easier to solve using EGF's rather than GF's.

Example 40:

$$g_0 = 0\,,\ 3g_n = ng_{n-1} + 2n! \quad (n \geqslant 1)$$

Step 1: $n = 0$, $LHS = 0$, $RHS = 2$. So

$$3g_n = ng_{n-1} + 2n! - 2[n = 0]$$

Step 2: Multiply by $z^n/n!$ and sum

$$\frac{3g_n z^n}{n!} = \frac{g_{n-1}z^n}{(n-1)!} + 2z^n - 2[n=0]\frac{z^n}{n!}$$

Sum over n for $n \geqslant 0$.

$$\begin{aligned} 3\hat{G}(z) &= z\hat{G}(z) + \frac{2}{1-z} - 2 \\ \Rightarrow \hat{G}(z) &= \frac{2}{(1-z)(3-z)} - \frac{2}{3-z} \\ &= \frac{1}{1-z} - \frac{1}{3-z} - \frac{2}{3-z} \\ &= \frac{1}{1-z} - \frac{1}{1-z/3} = \sum z^n - \sum \frac{z^n}{3^n} \end{aligned}$$

So $g_n = n! - n!/3^n$.

Binomial Convolution
Let $\hat{F}(z)$ and $\hat{G}(z)$ be EGF's of $(f_0, f_1, \ldots)$ and $(g_0, g_1, \ldots)$.

Let

$$\begin{aligned} \hat{H}(z) = \hat{F}(z)\hat{G}(z) &= \left(\sum_{i\geqslant 0} \frac{f_i z^i}{i!}\right)\left(\sum_{j\geqslant 0} \frac{g_j z^j}{j!}\right) \\ &= \sum_{n\geqslant 0}\left(\sum_k \frac{f_k g_{n-k}}{k!(n-k)!}\right) z^n = \sum_{n\geqslant 0} \frac{h_n}{n!}\, z^n \end{aligned}$$

where

$$h_n = \sum_k \frac{n!}{k!(n-k)!} f_k g_{n-k} = \sum_k \binom{n}{k} f_k g_{n-k}.$$

We call the sequence (h_n) defined by this formula the *binomial convolution* of the sequences (f_n) and (g_n). Thus, binomial convolution of sequences corresponds to multiplication of their EGF's.

Example 41: Bernoulli numbers.

Definition. The *Bernoulli numbers* B_n $(n \geqslant 0)$ are defined by

$$\frac{z}{e^z - 1} = \sum_{n \geqslant 0} B_n \frac{z^n}{n!}. \tag{32}$$

The denominator $e^z - 1$ starts as $z + \frac{1}{2}z^2 + \cdots$, which is why one puts a z in the numerator (otherwise the sum on the right would have to start at $n = -1$).

Exercise. Use binomial convolution to show $\displaystyle\sum_{j=0}^{m} \binom{m+1}{j} B_j = [m = 0]$.

Exercise. Show that $B_n = 0$ if $n > 2$ is odd. Hint: let $f(z)$ denote the function defined in (32) and consider $f(z) - f(-z)$.

Here are some values for the Bernoulli numbers.

n	0	1	2	3	4	5	6	7	8	9	10
B_n	1	$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$	0	$-\frac{1}{30}$	0	$\frac{5}{66}$

Definition. The *Bernoulli polynomials* $B_n(x)$ $(n \geqslant 0)$ are defined by

$$\frac{z\, e^{xz}}{e^z - 1} = \sum_{n \geqslant 0} B_n(x) \frac{z^n}{n!}. \tag{33}$$

It is clear that

$$B_n(0) = B_n.$$

We will now show that $B_n(x)$ is indeed a polynomial, and we will express these polynomials in terms of the Bernoulli numbers. We have

$$\sum_{n \geqslant 0} B_n(x) \frac{z^n}{n!} \overset{(33)}{=} \frac{z}{e^z - 1} e^{xz} = \left(\sum_{k \geqslant 0} B_k \frac{z^k}{k!} \right) \left(\sum_{k \geqslant 0} x^k \frac{z^k}{k!} \right)$$

where the last equality is by (32) and the exponential series. By binomial convolution we find

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k\, x^{n-k}.$$

Proposition 18

$$\sum_{k=0}^{n} k^{m-1} = \frac{B_m(n+1) - B_m}{m}.$$

Proof. We have

$$\sum_{m\geqslant 0}\Big(B_m(x+1)-B_m(x)\Big)\,\frac{z^m}{m!} \overset{(33)}{=} \frac{z\,e^{(x+1)z}}{e^z-1}-\frac{z\,e^{xz}}{e^z-1}$$

$$=\frac{z\,e^{xz}(e^z-1)}{e^z-1}=z\,e^{xz}=\sum_{m\geqslant 0}x^m\,\frac{z^{m+1}}{m!}=\sum_{m\geqslant 1}x^{m-1}\,\frac{z^m}{(m-1)!}.$$

Taking coefficients of z^m on both sides gives

$$B_m(x+1)-B_m(x)=m\,x^{m-1}.$$

Write $k=x$ and sum over $k\in\{0,1,\ldots,n\}$:

$$\begin{aligned} m\sum_{k=0}^{n}k^{m-1} &= \sum_{k=0}^{n}\Big(B_m(k+1)-B_m(k)\Big)\\ &= B_m(n+1)-B_m(0)=B_m(n+1)-B_m. \end{aligned}$$

□

Example 42: Bell numbers.

The *Bell number* $b(n)$ is the number of partitions of X_n. So we have

$$b(n)=\sum_{k=0}^{n}\left\{ {n \atop k} \right\}.$$

We have the following recursion:

$$\begin{aligned} b(n+1) &= \sum_{k\geqslant 0}\begin{pmatrix}\text{number of partitions of } X_{n+1} \text{ such}\\ \text{that } n+1 \text{ is in a part of size } k+1\end{pmatrix}\\ &= \sum_{k=0}^{n}\binom{n}{k}b(n-k), \quad (n\geqslant 0). \end{aligned}$$

So if we put

$$B(z)=\sum_{n\geqslant 0}b(n)\frac{z^n}{n!}=\text{exponential GF of the Bell numbers}$$

then we find, on differentiating

$$\begin{aligned} B'(z) &= \sum_{n\geqslant 1}b(n)\frac{z^{n-1}}{(n-1)!}=\sum_{n\geqslant 0}b(n+1)\frac{z^n}{n!}\\ &= \sum_{n\geqslant 0}\sum_{k=0}^{n}\binom{n}{k}b(k)\frac{z^n}{n!}=\sum_{n\geqslant 0}\sum_{k=0}^{n}\frac{b(k)}{k!}\frac{1}{(n-k)!}z^n\\ &= \left(\sum_{k\geqslant 0}\frac{b(k)}{k!}z^k\right)\left(\sum_{\ell\geqslant 0}\frac{z^\ell}{\ell!}\right)\\ &= B(z)e^z. \end{aligned}$$

Therefore,

$$\frac{d}{dz}\log B(z) = \frac{B'(z)}{B(z)} = e^z, \quad \log B(z) = e^z + c,$$

and

$$B(z) = e^{e^z + c}$$

for some constant c. Since $e^{1+c} = B(0) = b(0) = 1$ we must have $c = -1$ so

$$B(z) = e^{e^z - 1}.$$

We thus find a closed formula for the exponential generating function of the Bell numbers, but a closed formula for the Bell numbers is not possible.

Exercise. Prove $b(n) = e^{-1}\sum_{k\geqslant 0}\frac{k^n}{k!}$. (Ignore convergence questions.)

Example 43: Higher Derangements.

In section 4.5 we considered derangements: a derangement of $X_n = \{1,\ldots,n\}$ is a permutation of X_n without fixed points. We found the following formula for the number $D(n)$ of derangements of X_n:

$$D(n) = n!\sum_{\ell=0}^{n}\frac{(-1)^\ell}{\ell!}.$$

A k-derangement of X_n is a permutation g of X_n all of whose orbits have more than k elements (recall that a g-orbit is a set of the form $\{g^m(x) \mid m \in \mathbb{Z}\}$ where $x \in X_n$). Let $D(n,k)$ denote the number of k-derangements of X_n.

So a 1-derangement is just a derangement, and $D(n,1) = D(n)$.

We will find a formula for the exponential generating function

$$f(x) := \sum_{n\geqslant 0}\frac{D(n,k)}{n!}x^n.$$

Let $E(n+1,k,p)$ denote the set of k-derangements in S_{n+1} such that $n+1$ is in a $(p+1)$-cycle. So

$$D(n+1,k) = \sum_{p\geqslant k}\#E(n+1,k,p). \tag{34}$$

We compute $\#E(n+1,k,p)$. One chooses an element g of $E(n+1,k,p)$ in three steps.

Step 1. Choose the orbit of $n+1$, that is, $\{g^\ell(n+1) : \ell \in \mathbb{Z}\}$. There are $\binom{n}{p}$ choices because the orbit has $p+1$ elements.

Step 2. Choose the cycle of $n+1$. There are $p!$ choices.

Step 3. Finish g. There are $(n+1)-(p+1) = n-p$ elements left to permute so the number of choices is $D(n-p,k)$.

We are dealing with a homogeneous tree so

$$\#E(n+1,k,p) = \binom{n}{p}\, p!\, D(n-p,k) = \frac{n!\, D(n-p,k)}{(n-p)!}$$

and by (34)

$$D(n+1,k) = \sum_{p=k}^{n} \frac{n!\, D(n-p,k)}{(n-p)!}. \tag{35}$$

We find

$$\begin{aligned}
f'(x) &= \sum_{n\geqslant 1} \frac{D(n,k)\, x^{n-1}}{(n-1)!} = \sum_{n\geqslant 0} \frac{D(n+1,k)\, x^n}{n!} \\
&= \sum_{n\geqslant k} \frac{D(n+1,k)\, x^n}{n!} \qquad \text{because the other terms are zero} \\
&= \sum_{n\geqslant k} \frac{x^n}{n!} \left(\sum_{p=k}^{n} \frac{n!\, D(n-p,k)}{(n-p)!} \right) \qquad \text{by (35)} \\
&= \sum_{p\geqslant k} \sum_{m\geqslant 0} \frac{x^{m+p}\, D(m,k)}{m!} \\
&= \left(\sum_{p\geqslant k} x^p \right) \left(\sum_{m\geqslant 0} \frac{D(m,k)\, x^m}{m!} \right) = \frac{x^k}{1-x}\, f(x).
\end{aligned}$$

In section 6.5 we have learned how to solve a differential equation like the one here, $f'(x) = x^k\,(1-x)^{-1}\, f(x)$. We find

$$\begin{aligned}
\frac{f'}{f} &= \frac{x^k}{1-x}, \\
(\log f)' &= x^k + x^{k+1} + \cdots = \frac{1}{1-x} - \left(1 + x + x^2 + \cdots + x^{k-1}\right), \\
\log f &= -\log(1-x) - \left(x + \frac{x}{2} + \frac{x^3}{3} + \cdots + \frac{x^k}{k}\right) + c \qquad (\text{and } c = 0), \\
f(x) &= \frac{e^{-(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^k}{k})}}{1-x}.
\end{aligned}$$

A closed formula for $D(n,k)$ is not possible.

Exercise. Show that

$$\lim_{n\to\infty} \frac{D(n,k)}{n!}$$

exists and compute it. (In this exercise, you may use, without proving it, that there are $a_0, a_1, \ldots \in \mathbb{C}$ such that

$$e^{-(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^k}{k})} = \sum_{n\geqslant 0} a_n x^n$$

for any complex number x, in the sense that the right hand side converges to the left hand side. By the way, if you know a little complex function theory then you can prove this by observing that the left hand side is a holomorphic function in x on the complex plane.)

6.7 Generating functions in more variables

Generating functions in more variables exist too.

Example 44: Define e^x to be the formal power series

$$e^x = \sum_{n\geqslant 0} \frac{x^n}{n!}.$$

Our aim is to prove

$$e^x\, e^y = e^{x+y}.$$

We don't allow ourselves to use any similar looking result from analysis. Since the assertion is about formal power series, we don't need to consider convergence questions.

The proof uses the binomial theorem 6 and goes as follows:

$$\begin{aligned} e^{x+y} &= \sum_{n\geqslant 0} \frac{(x+y)^n}{n!} = \sum_{n\geqslant 0}\sum_{k=0}^{n} \binom{n}{k} \frac{x^k y^{n-k}}{n!} \\ &= \sum_{n\geqslant 0}\sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} \frac{x^k\, y^{n-k}}{n!} = \sum_{n\geqslant 0}\sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \\ &= \left(\sum_{k\geqslant 0} \frac{x^k}{k!}\right)\left(\sum_{\ell\geqslant 0} \frac{y^\ell}{\ell!}\right) = e^x\, e^y. \end{aligned}$$

Exercise. (a) Prove: $\displaystyle\sum_{k,\ell\geqslant 0} \binom{k+\ell}{k} x^k\, y^\ell = \frac{1}{1-x-y}.$

(b) Put $a_n := \displaystyle\sum_{m=0}^{n} \binom{n-m}{m}$. Use (a) to compute $A(t) = \sum_{n\geqslant 0} a_n t^n$.

Exercise. Prove

$$\sum_{n_1,\ldots,n_k\geqslant 0} \min(n_1,\ldots,n_k)\, x_1^{n_1}\cdots x_k^{n_k} = \frac{x_1\cdots x_k}{(1-x_1)\cdots(1-x_k)(1-x_1\cdots x_k)}.$$

7 Discrete Probability

7.1 Sample Spaces and Random Variables

A discrete sample space is a set Ω together with a function $P : \Omega \to [0,1]$ (the probability function) such that $\sum_{\omega\in\Omega} P(\omega) = 1$. In most examples such a set is finite or countable; indeed, from the fact that $\sum_{\omega\in\Omega} P(\omega)$ converges to a finite sum, it follows that $P(\omega) = 0$ for all but countably many ω (Exercise: prove this. It's not hard — for each n, for how many ω can we have $P(\omega) \geqslant 1/n$?)

The space Ω is *uniform* if P is constant.

A subset A of Ω is called an *event*. We define

$$P(A) = \sum_{\omega\in A} P(\omega).$$

Example 45: $\Omega = \{1,2,3,4,5,6\}$, $P(\omega) = 1/6$ for all $\omega \in \Omega$ (for example throwing a dice)

$$A = \{5,6\},\ P(A) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

If $A, B \subseteq \Omega$ and $A \cap B = \varnothing$ then $P(A \cup B) = P(A) + P(B)$. More generally,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

(Exercise).

A random variable (RV), X, is a function defined on a sample space Ω. Usually, $X : \Omega \to \mathbb{R}$. For $x \in \text{Range}(X)$ we have

$$P(X = x) = \sum_{\substack{\omega\in\Omega \\ X(\omega)=x}} P(\omega)$$

Example 46: $\Omega = \{(i,j) \mid i,j \in [1\ldots6]\}$, $P(\omega) = \frac{1}{36}$ for all $\omega \in \Omega$ (throw dice twice). Then

$$\begin{aligned} S_1 : (i,j) &\to i \\ S_2 : (i,j) &\to j \\ S : (i,j) &\to i+j \end{aligned}$$

are all random variables defined on Ω and $S = S_1 + S_2$.

x	2	3	4	5	6	7	8	9	10	11	12
$P(S = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$\text{Range}(S) = [2\ldots12]$

Two random variables $X : \Omega \to T_1$ and $Y : \Omega \to T_2$ are called *independent* if

$$P(X = x \text{ and } Y = y) = P(X = x)P(Y = y)$$

for all $x \in T_1$ and $y \in T_2$. So in the above example S_1 and S_2 are independent but S_1 and $S_1 + S_2$ are not independent.

We define the *mean* or *expected value* $E(X)$ of the RV X to be

$$\sum_{x \in X(\Omega)} xP(X = x) = \sum_{\omega \in \Omega} X(\omega)P(\omega)$$

For example, for $S : (i, j) \to i + j$ as above

$$E(S) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + \ldots + 12 \times \frac{1}{36} = 7$$

In general,

$$E(X + Y) = E(X) + E(Y)$$
$$E(S_1) = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{7}{2}.$$

So

$$E(S) = E(S_1) + E(S_2) = \frac{7}{2} + \frac{7}{2} = 7.$$

What about $E(XY)$? If X and Y are independent, then we claim that $E(XY) = E(X)E(Y)$. In the following proof, we use the fact that Ω is partitioned into disjoint sets $\{\omega \in \Omega : X(\omega) = x, Y(\omega) = y\}$ as x and y vary in $X(\Omega)$ and $Y(\Omega)$.

$$\begin{aligned}
E(XY) &= \sum_{\omega \in \Omega} X(\omega)Y(\omega)P(\omega) \\
&= \sum_{x \in X(\Omega),\ y \in Y(\Omega)} \left(\sum_{X(\omega)=x,\ Y(\omega)=y} xyP(\omega) \right) \\
&= \sum_{x \in X(\Omega),\ y \in Y(\Omega)} xyP(X = x, Y = y) \\
&= \sum_{x \in X(\Omega),\ y \in Y(\Omega)} xyP(X = x)P(Y = y) \\
&= \sum_{x \in X(\Omega)} xP(X = x) \sum_{y \in Y(\Omega)} yP(Y = y) \\
&= E(X)E(Y).
\end{aligned}$$

For example $E(S_1S_2) = 49/4$. We often use μ, or $\mu(X)$, to denote the expected value $E(X)$.

The *variance* $V(X)$ of X is defined to be $E((X - \mu)^2)$; it is a measure of the deviation of X from the mean (of the spread of distribution). This has nicer mathematical properties than other possibilities such as $E(|X - \mu|)$

The *standard deviation* $\sigma(X)$ is $\sqrt{V(X)}$

$$\begin{aligned} V(X) &= E((X-\mu)^2) \\ &= E(X^2 - 2X\mu + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - \mu^2 \\ &= E(X^2) - E(X)^2 \end{aligned}$$

Let X and Y be two independent RVs on Ω.

$$\begin{aligned} V(X+Y) &= E((X+Y)^2) - E(X+Y)^2 \\ &= E(X^2) + E(Y^2) + 2E(XY) - E(X)^2 - E(Y)^2 - 2E(X)E(Y) \end{aligned}$$

But

$$E(XY) = E(X)E(Y)$$

by independence. So

$$V(X+Y) = V(X) + V(Y).$$

Note that

$$V(X+X) = V(2X) = 4V(X)$$

so we need independence. In the dice example

$$\begin{aligned} V(S_1) &= \frac{1^2 + 2^2 + \ldots + 6^2}{6} - \frac{49}{4} = \frac{35}{12}, \\ V(S) = V(S_1 + S_2) &= \frac{35}{6}. \end{aligned}$$

Choose any $\alpha \geqslant 0$,

$$\begin{aligned} V(X) &= \sum_{\omega\in\Omega} (X(\omega) - \mu)^2 P(\omega) \\ &\geqslant \sum_{\substack{\omega\in\Omega \\ (X(\omega)-\mu)^2 \geqslant \alpha}} (X(\omega) - \mu)^2 P(\omega) \\ &\geqslant \sum_{\substack{\omega\in\Omega \\ (X(\omega)-\mu)^2 \geqslant \alpha}} \alpha P(\omega) \\ &= \alpha P((X-\mu)^2 \geqslant \alpha) \end{aligned}$$

Define c by

$$\alpha = c^2\sigma^2 \quad \text{(where } \sigma = \sqrt{V} \text{ is the standard deviation of } X\text{)}$$

Then provided $\alpha > 0$,

$$\begin{aligned} P((X-\mu)^2 \geqslant \alpha) &\leqslant \frac{V(X)}{\alpha} \\ \Rightarrow P(|X-\mu| \geqslant c\sigma) &\leqslant \frac{1}{c^2} \end{aligned} \tag{36}$$

So the probability that the difference from the mean is greater than or equal to c times the standard deviation, is less than or equal to $1/c^2$. This is true for any RV on any sample space. In specific cases, we get a stronger result. For example, in the normal distribution

$$P(|X - \mu| \geqslant 2\sigma) \sim 0.05$$

We next discuss the notion of multiple *independent instances* of a single random variable X. This models the idea of repeated independent trials, e.g. n independent dice throws. Suppose Ω_1 is a sample space, with probability measure $P_1 : \Omega_1 \to [0,1]$. Define $\Omega_n = \Omega_1 \times \cdots \times \Omega_1$ (n times). We define a probability measure $P_n : \Omega_n \to [0,1]$ by

$$P_n(\omega_1, \ldots, \omega_n) = P_1(\omega_1) \times \cdots \times P_1(\omega_n).$$

Exercise: Show that this is a probability measure, i.e. that

$$\sum_{(\omega_1,\ldots,\omega_n)\in\Omega_n} P_n(\omega_1 \ldots, \omega_n) = 1.$$

More generally, show that if $A_i \subset \Omega_1$ for $i = 1, \ldots, n$ then

$$P_n(A_1 \times \cdots \times A_n) = P_1(A_1) \times \cdots \times P_1(A_n).$$

We get our "n independent instances of X", which we now call $X_1, \ldots, X_n$, by defining

$$X_i(\omega_1, \ldots, \omega_n) = X(\omega_i).$$

Exercise: Show that the random variables X_i and X_j are independent for $i \neq j$.

It is more or less obvious that

$$P_n(X_i = x) = P_1(X = x);$$

for the set

$$\{(\omega_1, \ldots, \omega_n) : X_i(\omega_1, \ldots, \omega_n) = x\}$$

is equal to

$$\Omega_1 \times \cdots \times \Omega_1 \times \{\omega_i : X(\omega_i) = x\} \times \Omega_1 \times \cdots \times \Omega_1.$$

Given n independent instances of X, say $X_1, X_2, \ldots X_n$, let $S = X_1 + \cdots + X_n$. Then

$$\begin{aligned} E(S) &= nE(X) \\ &= n\mu \\ V(S) &= nV(X) \\ \sigma(S) &= \sqrt{n}\sigma \\ E(S/n) &= \mu \\ \sigma(S/n) &= \frac{\sigma}{\sqrt{n}} \end{aligned}$$

The spread of averages of n instances is smaller than the spread of single instances.

$$P\left(\left|\frac{S}{n}-\mu\right| \geqslant \frac{c\sigma}{\sqrt{n}}\right) \leqslant \frac{1}{c^2}\,.$$

Or putting $d = c/\sqrt{n}$,

$$P\left(\left|\frac{S}{n}-\mu\right| \geqslant d\sigma\right) \leqslant \frac{1}{nd^2}\,.$$

7.2 Probability Generating Functions (PGF's)

Let Ω be a sample space and X a RV taking values in $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Let

$$g_k = P(X = k)\,, \text{ for } k \geqslant 0$$

Then the PGF of the RV X is just the generating function of the sequence (g_k).

$$G_X(z) \;=\; \sum_{k\geqslant 0} g_k z^k$$

Since

$$\sum_{k\geqslant 0} g_k = 1$$

(for the subsets $\{X = k\}$ partition the sample space Ω), we must have $G_X(1) = 1$. And so $G_X(1)$ must be a convergent series. Moreover,

$$E(X) = \sum_k P(X = k)k = G'_X(1)\,, \quad \text{where } G'_X \text{ is the derivative of } G_X$$

This series is not always convergent, that is, $E(X)$ can be infinite:

Example 47: we toss a coin until we get tails, and if we get k heads before the tail, then you pay me 2^k pennies. The sample space Ω is

$$\Omega = \{t, ht, hht, hhht, \dots,\}.$$

Denote $h\cdots ht$ (k heads followed by a tail) by $h^k t$. Provided the coin is fair, heads and tails each have probability 1/2; we assume the outcomes of successive tosses are independent, so the probability measure $P : \Omega \to [0, 1]$ is given by $P(h^k t) = (1/2)^{k+1}$. The random variable X is $X(h^k t) = 2^k$, so

$$E(X) = \sum_{k\geqslant 0} \frac{2^k}{2^{k+1}} = \sum_{k\geqslant 0} \frac{1}{2} = \infty$$

For any random variable X,

$$\begin{aligned} E(X^2) &= \sum_{k\geqslant 0} k^2 P(X = k) \\ &= \sum_{k\geqslant 0} (k(k-1) + k)P(X = k) \\ &= G''_X(1) + G'_X(1) \end{aligned}$$

$$V(X) = E(X^2) - E(X)^2 = G''_X(1) + G'_X(1) - G'_X(1)^2.$$

Example 48: A random variable $U_n : \Omega \to [0, \cdots, n-1] \subset \mathbb{N}_0$ with uniform probability distribution: $U_n(\omega) = 0, 1, \ldots, n-1$ each with probability $1/n$. The PGF is

$$G_{U_n}(z) = \frac{1 + z + z^2 + \ldots + z^{n-1}}{n} = \frac{(1 - z^n)}{n(1 - z)}, \text{ if } z \neq 1$$

which is unfortunate, since to find $E(U_n)$ and $V(U_n)$ we want to be able to evaluate G'_{U_n} and G''_{U_n} when $z = 1$! To get round this, use Taylor's Theorem around $z = 1$:

$$G_{U_n}(1 + z) = G_{U_n}(1) + G'_{U_n}(1)z + G''_{U_n}(1)z^2/2 + \ldots$$

Get

$$\begin{aligned} G_{U_n}(1 + z) &= \frac{1 - (1 + z)^n}{-nz} = \frac{(1 + z)^n - 1}{nz} \\ &= \frac{1}{nz}\left(nz + \binom{n}{2} z^2 + \binom{n}{3} z^3 + \ldots\right) \\ &= 1 + \frac{(n-1)}{2} z + \frac{(n-1)(n-2)}{6} z^2 + \ldots \end{aligned}$$

So

$$\begin{aligned} G_{U_n}(1) &= 1 \\ G'_{U_n}(1) &= \frac{n-1}{2} \\ G''_{U_n}(1) &= \frac{(n-1)(n-2)}{3} \end{aligned}$$

$$\begin{aligned} E(U_n) &= G'_{U_n}(1) = \frac{n-1}{2} \\ V(U_n) &= G''_{U_n}(1) + G'_{U_n}(1) - G'_{U_n}(1)^2 = \frac{n^2 - 1}{12} \end{aligned}$$

For example, if $n = 6$,

$$V(U_6) = \frac{35}{12}$$

Let X, Y be independent RVs on Ω, $X, Y : \Omega \to \mathbb{N}_0$. Let $g_k = P(X = k)$, $h_k = P(Y = k)$. Consider $X + Y$:

$$\begin{aligned} P(X + Y = n) &= \sum_k P(X = k \text{ and } Y = n - k) \\ &= \sum_k P(X = k)P(Y = n - k), \text{ by independence} \\ &= \sum_k g_k h_{n-k} \end{aligned}$$

So PGF of $X+Y$ is

$$G_{X+Y}(z) = \sum_{n\geqslant 0} \Big(\sum_k g_k h_{n-k} \Big) z^n = G_X(z) G_Y(z).$$

In particular, if S is the sum of n independent instances of X, then

$$G_S(z) = (G_X(z))^n$$

7.3 Tossing Coins

From now we write $X(z)$ instead of $G_X(z)$.

A single coin toss has two outcomes $h,\ t$ with probabilities p, q with $p+q=1$. A fair coin has $p=q=1/2$.

1. Basic example. $\Omega = \{h, t\}$ $P(h) = p$ and $P(t) = q$

Define an RV H on Ω by $H(h) = 1\,,\ H(t) = 0$. So PGF $H(z) = q + pz$. $P(H=0) = q$, $P(H=1) = p$. When we toss the coin n times, H_n, the sum of n instances of H, is the total number of h. The PGF of H_n is

$$H_n(z) \quad = \quad H(z)^n = (q+pz)^n = \sum_{k\geqslant 0} \binom{n}{k} p^k q^{n-k} z^k.$$

We can mine this for a great deal of information:

$$\begin{aligned}
P(H_n = k) &= \binom{n}{k} p^k q^{n-k}, \\
E(H) &= H'(1), \\
&= p, \\
E(H_n) &= nE(H) = np, \\
V(H) &= H''(1) + H'(1) - H'(1)^2 = 0 + p - p^2 = p(1-p) = pq, \\
V(H_n) &= nV(H) = npq, \\
\sigma(H_n) &= \sqrt{npq}.
\end{aligned}$$

2. Suppose we toss the coin until we get h then stop. $\Omega = \{h, th, tth, ttth, \ldots\} = \{t^k h \mid k \in \mathbb{N}_0\}$ where $t^k h$ means t k times, then h.

$$P(t^k h) = q^k p$$

$$\sum_{k\geqslant 0} q^k p = \frac{p}{1-q} = \frac{p}{p} = 1$$

So we do have a sample-space, that is, $\sum_{\omega\in\Omega} P(\omega) = 1$.

Definition 7: Define a RV F on Ω by $F(t^k h) = k+1 =$ total number of tosses to get h.

Then

$$F(z) = \sum_{k \geqslant 1} q^{k-1} p z^k = \frac{pz}{1-qz}.$$

Again, F_n = sum of n independent instances of F. This represents the number of tosses needed to get h n times.

The PGF of F_n is

$$\left(\frac{pz}{1-qz}\right)^n.$$

So

$$\begin{aligned}
F'(z) &= \frac{(1-qz)p + pzq}{(1-qz)^2} = \frac{p}{(1-qz)^2} \\
F'(1) &= \frac{p}{(1-q)^2} = \frac{1}{p}, \\
E(F) &= \frac{1}{p}, \\
E(F_n) &= \frac{n}{p}, \\
F''(z) &= \frac{2pq}{(1-qz)^3} \\
F''(1) &= \frac{2q}{p^2}, \\
V(F) &= \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}, \\
V(F_n) &= \frac{nq}{p^2}
\end{aligned}$$

3. Now we keep tossing the coin until we get h twice in a row, hh. The sample space is $\{hh, thh, hthh, tthh, \ldots\}$ with probabilities $\{p^2, qp^2, qp^3, q^2p^2, \ldots\}$. Let G on Ω be an RV equal to total number of tosses as in (2).

$G(z) = p^2z^2 + qp^2z^3 + (qp^3 + q^2p^2)z^4 + \ldots$ We get $G(z)$ by substituting qz for t and pz for h in each element of Ω and summing. An element of Ω consists of an arbitrary non-empty string in t and ht followed by hh.

Strings of length 1 in t and ht are t, ht. Length 2: $tt, tht, htt, htht$. Adding these up, this could be written as $(t+ht)^2$. Adding up strings of length 3 in t and ht gives $(t+ht)^3$. Adding up strings of length k in t and ht gives $(t+ht)^k$. So sum of elements of Ω can be written as

$$\sum_{k \geqslant 0} (t+ht)^k hh$$

Now substitute qz for t and pz for h.

So PGF

$$G(z) = \sum_{k\geqslant 0}(qz+pqz^2)^k p^2z^2 = \frac{p^2z^2}{1-qz-pqz^2}G(1)$$
$$= \frac{p^2}{1-q-pq} = \frac{p^2}{p-pq} = \frac{p^2}{p(1-q)} = 1\,.$$

So

$$\sum_{\omega\in\Omega} P(\omega) = 1$$

and Ω really is a sample space. That is, sequence hh occurs with probability 1 if you keep tossing the coin.

$$G'(z) = \frac{2(1-qz-pqz^2)p^2z + p^2z^2(q+2pqz)}{(1-qz-pqz^2)^2}$$

$$\text{expected number of tosses} = G'(1) = \frac{2p^4+p^2(q+2pq)}{p^4}$$
$$= \frac{2p^4+p^2(1-2p^2+p)}{p^4} = \frac{1}{p}+\frac{1}{p^2}.$$

For example, if $p=\frac{1}{2}$, $G'(1)=6$.

4. Toss a coin until we get hht.

A sequence ending with first occurence of hht is an arbitrary string of $k(\geqslant 0)$ t or ht followed by string of $l(\geqslant 0)$ h, followed by hht.

Using the same idea as in the last example,

$$\sum_{\Omega}\omega \;=\; \sum_{k\geqslant 0}\sum_{l\geqslant 0}(t+ht)^k h^l hht = \Big(\sum_{k\geqslant 0}(t+ht)^k\Big)\Big(\sum_{l\geqslant 0}h^l\Big)hht$$

For generating function $t\to qz$, $h\to pz$.

$$\sum_{k\geqslant 0}(qz+pqz^2)^k\Big(\sum_{\ell\geqslant 0}(pz)^\ell p^2qz^3\Big) = \frac{p^2qz^3}{(1-qz-pqz^2)(1-pz)}$$

In case $p=1/2$

$$\begin{aligned} G(z) &= \frac{z^3}{z^3-8z+8} \\ G'(z) &= \frac{(z^3-8z+8)3z^2 - z^3(3z^2-8)}{(z^3-8z+8)^2} \\ G'(1) &= 8 \end{aligned}$$

the expected number of tosses.

If we did thh instead, element of Ω is $l\geqslant 0$ $h's$ followed by $k\geqslant 0$ t or ht followed by thh. So the generating function is the same and so is the expected number.

5. Consider the following game between A and B. We toss a fair coin ($p = 1/2$) until either hht occurs or thh. A wins if hht comes first, B if thh first.

A wins with pattern $h^k t$ ($k \geqslant 2$) only (that is, A wins only if first two tosses are h, with probability $1/4$). B wins with $(t + ht)^k hh \quad (k \geqslant 1)$. The generating function for the game is the GF of $h^k t$ plus the GF of $(t+ht)^k hh$.

$$G(z) = \frac{z}{4(2-z)} + \frac{(z/2 + z^2/4)\, z^2/4}{1 - z/2 - z^2/4}.$$

Putting $z = 1$ gives

$$\underset{(A \text{ wins})}{\tfrac{1}{4}} \quad + \quad \underset{(B \text{ wins})}{\tfrac{3}{4}}$$

So thh is "better combination than hht".

6. A new game: now suppose that A wins with hht and B wins with htt.

Winning pattern for A is $t^k (ht)^l h^m hht$, $k, l, m \geqslant 0$. B wins with $t^k\ (ht)^l htt$, $k, l \geqslant 0$

The GF is

$$\underset{(A \text{ wins})}{\frac{z^3}{8(1-z/2)(1-z^2/4)(1-z/2)}} \quad + \quad \underset{(B \text{ wins})}{\frac{z^3}{8(1-z/2)(1-z^2/4)}}$$

Putting $z = 1$ gives

$$\frac{2}{3} + \frac{1}{3}.$$

A wins with probability $2/3$.

So hht is better than htt. We have $thh > hht > htt$.

But thh and htt are equally likely to come first (by symmetry). So the relationship 'better than' is *not* transitive.