Chapter 1

Counting Functions and Subsets

The goal of this short chapter is to introduce the very important numbers $\binom{n}{r}$, called the **binomial coefficients**, and to prove the Binomial Theorem. The binomial coefficients have many uses in mathematics beyond the binomial theorem — in combinatorics, probability, statistics; they are even linked to the Riemann Hypothesis. In this course they will, for example, be used in one of the proofs of Fermat's Little Theorem. It turns out that to count subsets, it's convenient to count functions first.

Notation. Given finite sets X and Y, let

- $Fun\,(X,Y)$ be the set of **all functions** from X to Y;
- $Inj\,(X,Y)$ be the set of **all injective functions** from X to Y;
- $Bij\,(X,Y)$ be the set of **all bijective functions** from X to Y;

Proposition 1.1 (Counting functions)**.** If $|X| = m$, $|Y| = n$, then

- $|Fun(X,Y)| = n^m$;
- $|Inj(X,Y)| = \dfrac{n!}{(n-m)!}$ if $m \leq n$, 0 if $m > n$;
- $|Bij(X,Y)| = n!$ if $m = n$, 0 if $m \neq n$.

Proof. To make notation easier we assume $X = \mathbb{N}_m$. We use induction in m to prove that, if $|Y| = n \geq m$, there are $\dfrac{n!}{(n-m)!}$ injective functions from $\mathbb{N}_m$ to Y.

The base case is $m = 1$. For each element $y \in Y$ there is one function from $\mathbb{N}_1 = \{1\}$ to Y given by $1 \mapsto y$, and all these are injective functions. So there are n injective functions. We have $\frac{n!}{(n-1)!} = n$ so the statement is true for $m = 1$.

The inductive step: assume that there are $\frac{n!}{(n-m)!}$ injective functions from $\mathbb{N}_m$ to Y. *Instead of counting injective functions* $f\colon \mathbb{N}_{m+1} \to Y$ *we will count their* **restrictions** $f|_{\mathbb{N}_m}$ which are also injective, so there are $\frac{n!}{(n-m)!}$ of them. Many functions f with domain $\mathbb{N}_{m+1}$ have the same restriction onto $\mathbb{N}_m$: the value $f(m+1)$ can be any element of Y except $f(1)$, $f(2)$, ..., $f(m)$ to ensure injectivity, hence $n - m$ functions $f \in Inj(\mathbb{N}_{m+1}, Y)$ have the same restriction $f|_{\mathbb{N}_m} \in Inj(\mathbb{N}_m, Y)$.

Since every injective function from $\mathbb{N}_m$ to Y is the restriction of $(n-m)$ injective functions from $\mathbb{N}_{m+1}$ to Y, it follows that $|Inj(\mathbb{N}_{m+1}, Y)| = |Inj(\mathbb{N}_m, Y)| \times (n-m)$, giving

$$\frac{n!}{(n-m)!}(n-m) = \frac{n!}{(n-m-1)!(n-m)}(n-m) = \frac{n!}{(n-(m+1))!}.$$

By induction, the formula for $|Inj(X, Y)|$ is true for all m.

The formula $|Fun(X, Y)| = n^m$ is left to the students as an exercise. (Idea: n functions $f \in Fun(\mathbb{N}_{m+1}, Y)$ have the same restriction $f|_{\mathbb{N}_m} \in Fun(\mathbb{N}_m, Y)$, therefore $|Fun(\mathbb{N}_{m+1}, Y)| = |Fun(\mathbb{N}_m, Y)| \times n$.)

Finally, bijections from X to Y do not exist if $m \neq n$. If $m = n$, every injective function is surjective and bijective so $|Bij(X, Y)| = |Inj(X, Y)| = \frac{n!}{(n-n)!} = n!$ as $0! = 1$. □

Here is the first example where we count functions to count **subsets**.

Notation. If A is a set, $\mathcal{P}(A)$ is the **power set** of A. That is, $\mathcal{P}(A) = \{C \mid C \subseteq A\}$.

Proposition 1.2 (Counting all subsets)**.** If A is a finite set, then $|\mathcal{P}(A)| = 2^{|A|}$.

Proof. *Instead of counting subsets of* A*, we will count functions from* A *to* $\{0, 1\}$.

Given $C \subseteq A$, define $f\colon A \to \{0, 1\}$ by $f(x) = 1$ if $x \in C$, $f(x) = 0$ if $x \notin C$. Then each subset of A corresponds to a function in $Fun(A, \{0, 1\})$, and each $f \in Fun(A, \{0, 1\})$ corresponds to a unique set $C = \{x \in A : f(x) = 1\}$. Hence we have constructed a bijection between $\mathcal{P}(A)$ and $Fun(A, \{0, 1\})$. This implies that $|\mathcal{P}(A)| = |Fun(A, \{0, 1\})|$ which is $2^{|A|}$ by Proposition 1.1. □

We will now want to count subsets of a certain size r.

Notation. $\mathcal{P}_r(A)$ denote the set of all subsets of A containing *exactly* r elements:

$$\mathcal{P}_r(A) = \{C \in \mathcal{P}(A) : |C| = r\}.$$

Definition. The **binomial coefficient** $\binom{n}{r}$, read as "n choose r", is the cardinality $|\mathcal{P}_r(A)|$ for *any* set A of cardinality n.

Example. The set $A = \{a, b, c, d, e\}$ has the following 3-element subsets: $\{a, b, c\}$, $\{a, b, d\}$, $\{a, b, e\}$, $\{a, c, d\}$, $\{a, c, e\}$, $\{a, d, e\}$, $\{b, c, d\}$, $\{b, c, e\}$, $\{b, d, e\}$, $\{c, d, e\}$. So $\binom{5}{3} = 10$.

Also, $\mathcal{P}_0(A) = \{\varnothing\}$ and $\mathcal{P}_5(A) = \{A\}$. So $\binom{5}{0} = 1$ and $\binom{5}{5} = 1$.

Remark. Writing down all the r-element subsets of an n-element set takes too long so we will give two other ways of calculating $\binom{n}{r}$.

Proposition 1.3 (The Factorial Formula)**.** $\binom{n}{r} = \dfrac{n!}{r!(n-r)!}$ for all $n \geq 0$ and $0 \leq r \leq n$.

Proof. *Instead of counting r-element subsets of A we will count injective functions $f\colon \mathbb{N}_r \to A$.*

Each $f \in Inj(\mathbb{N}_r, A)$ can be written as

$$\mathbb{N}_r \xrightarrow{f} \mathrm{Im}f \subset A,$$

i.e. as a **bijection** from $\mathbb{N}_r$ onto the r-element subset $C = \mathrm{Im}f$ of A. The same r-element set C is the image of many injective functions: by Proposition 1.1, for a fixed $C \in \mathcal{P}_r(A)$ there are $r!$ bijections $f\colon \mathbb{N}_r \to C$. Hence $|Inj(\mathbb{N}_r, A)| = r! \times |\mathcal{P}_r(A)|$. Substituting, we obtain $\dfrac{n!}{(n-r)!} = r!\binom{n}{r}$ as claimed. □

Proposition 1.4 (The Inductive Formula)**.** For all $n \geq 1$ and all r, $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

Proof. The right-hand side is

$$\begin{aligned}\frac{(n-1)!}{(n-1-r)!\,r!} + \frac{(n-1)!}{(n-r)!(r-1)!} &= \frac{(n-1)!\,(n-r)}{(n-r)!\,r!} + \frac{(n-1)!\,r}{(n-r)!\,r!} \\ &= \frac{(n-1)!\,n}{(n-r)!\,r!} \\ &= \binom{n}{r}.\end{aligned}$$

□

Remark. A different way to prove the Inductive Formula would be to observe that there are $\binom{n-1}{r}$ r-element subsets of $\mathbb{N}_n$ which do not contain the element n, and there are $\binom{n-1}{r-1}$ r-element subsets of $\mathbb{N}_n$ which contain the element n (because the elements except n form an $(r-1)$-element subset of $\mathbb{N}_{n-1}$). Altogether, these are all r-element subsets of $\mathbb{N}_n$.

Remark (Pascal's Triangle)**.** The result of Proposition 1.4 is usually represented as an unending triangle where each term, apart from those at the end of the rows, are the sum of the two terms in the line above.

$$\begin{array}{ccccccccccccc}
 & & & & & & \binom{0}{0} & & & & & & \\
 & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & \\
 & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
 & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & \\
 & & & & & & \vdots & & & & & & \\
 & & \binom{n-1}{0} & \cdots\cdots & \binom{n-1}{r-1} & & \binom{n-1}{r} & & \cdots\cdots & \binom{n-1}{n-1} & & & \\
 & \binom{n}{0} & & \cdots\cdots\cdots\cdots & & \binom{n}{r} & & \cdots\cdots\cdots\cdots & & \binom{n}{n} & & & \\
 & & & & & & \vdots & & & & & &
\end{array}$$

The start of this is normally written as

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

We are ready to prove the famous

Theorem 1.5 (The Binomial Theorem)**.** For $a, b \in \mathbb{R}$, $n \geq 0$, we have $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$.

Remark. The right-hand side can also be written as $a^n + na^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \ldots + \binom{n}{n-2}a^2b^{n-2} + nab^{n-1} + b^n$. We use the convention that $x^0 = 1$ for all x.

Proof. First, we prove that for all $n \geq 0$, the statement $P(n)$: $(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$ is true.

Base case, $n = 0$, $P(0)$: $1 = \sum_{i=0}^{0} \binom{0}{0} x^0$ is true.

Inductive step: denote by $\operatorname{coef}_{x^i} f(x)$ the coefficient of x^i in the polynomial $f(x)$. Assume $P(k)$ true. Since $(1+x)^{k+1} = (1+x)^k + x(1+x)^k$,

$$\begin{aligned}
\operatorname{coef}_{x^i} (1+x)^{k+1} &= \operatorname{coef}_{x^i} (1+x)^k + \operatorname{coef}_{x^i} \left(x(1+x)^k\right) \\
&= \operatorname{coef}_{x^i} (1+x)^k + \operatorname{coef}_{x^{i-1}} (1+x)^k \\
&= \binom{k}{i} + \binom{k}{i-1} \quad \text{by } P(k) \\
&= \binom{k+1}{i}
\end{aligned}$$

by the Inductive Formula, Proposition 1.4. Thus $P(k+1)$ is true. By induction, $P(n)$ is true for all $n \geq 0$.

Now substitute $x = \dfrac{b}{a}$ to obtain $\left(1 + \dfrac{b}{a}\right)^n = \sum_{i=0}^{n} \binom{n}{i} \dfrac{b^i}{a^i}$. Multiply both sides by a^n to obtain $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$ as claimed. This does not work if $a = 0$, but then $(0+b)^n = 0 + 0 + \ldots + 0 + \binom{n}{n} b^n$ is trivially true. □

Remark (Advice for exam)**.** Theorem 1.5 is the Binomial Theorem, **not** Proposition 1.4 or Proposition 1.3.

Example.

- From the sixth line in Pascal's triangle we see

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5.$$

- The coefficient of a^7b^3 in $(2a+b)^{10}$ is

$$2^7 \binom{10}{3} = 2^7 \frac{10!}{3!7!} = 128 \times \frac{10 \times 9 \times 8}{3 \times 2} = 15360.$$

Remark (The sum and the alternating sum of the binomial coefficients)**.** The definition of $\binom{n}{r}$ and the Binomial Theorem imply two important facts about the binomial coefficients.

First, the **sum** of the binomial coefficients for a given n, i.e., the sum of the numbers in row n of Pascal's triangle, is

$$\sum_{r=0}^{n} \binom{n}{r} = 2^n.$$

Second, the **alternating sum** of the binomial coefficients for a given n, which means the sign-changing sum

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n} = \sum_{r=0}^{n} (-1)^r \binom{n}{r},$$

is zero.

Proofs of these two facts appear as questions in homework problem sheets.