# Chapter 10

# Binomial Coefficients

## 10.1 Basic properties

Recall that $\binom{n}{k}$ is the number of $k$-element subsets of an $n$-element set, and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{\prod_{i=0}^{k-1}(n-i)}{k!}.$$

The quantities $\binom{n}{k}$ are called *binomial coefficients* because of their role in the *Binomial Theorem*, apparently known to the 11th century Persian scholar Omar Khayyam. Before we state and prove the theorem let us consider some important identities that involve binomial coefficients. One that follows immediately from the algebraic definition is

$$\binom{n}{k} = \binom{n}{n-k}.$$

This also has a nice combinatorial interpretation: Choosing a $k$-element subset $B$ from an $n$-element set uniquely identifies the complement $A \setminus B$ of $B$ in $A$, which is an $(n-k)$-subset of $A$. This defines a bijection between $k$-element and $(n-k)$-element subsets of $A$, which implies the identity.

Another relation between binomial coefficients is called *Pascal's rule*, although it was known centuries before Pascal's time in the Middle East and India:

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

This can be easily proved algebraically:

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

Pascal's rule also has a combinatorial interpretation: $\binom{n+1}{k}$ is the number of $k$-element subsets of an $n$-element set $A$. Fix an element $a \in A$. A subset of $A$ either contains $a$ or it doesn't. $k$-element subsets of $A$ that do not contain $a$ are in fact $k$-element subsets of $A \setminus \{a\}$ and their number is $\binom{n}{k}$. $k$-element subsets of $A$ that do contain $a$ bijectively correspond to $(k-1)$-element subsets of $A \setminus \{a\}$, the number of which is $\binom{n}{k-1}$. The identity follows.

Another illuminating identity is the *Vandermonde convolution*:

$$\binom{m+n}{l} = \sum_{k=0}^{l} \binom{m}{k}\binom{n}{l-k}.$$

We only give a combinatorial argument for this one. We are counting the number of ways to choose an $l$-element subset of an $(m+n)$-element set $A$. Fix an $m$-element subset $B \subseteq A$. Any $l$-element subset $S$ of $A$ has $k$ elements from $B$ and $l-k$ elements from $A \setminus B$, for some $0 \leq k \leq l$. For a particular value of $k$, the number of $k$-element subsets of $B$ that can be part of $S$ is $\binom{m}{k}$ and the number of $(l-k)$-element subsets of $A \setminus B$ is $\binom{n}{l-k}$. We can now use the sum principle to sum over the possible values of $k$ and obtain the identity. An interesting special case is

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

It follows from the Vandermonde convolution by taking $l = m = n$ and remembering that $\binom{n}{k} = \binom{n}{n-k}$.

## 10.2 Binomial theorem

**Theorem 10.2.1.** *For $n \in \mathbb{N}$ and $x, y \in \mathbb{R}$,*

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

*Proof.* By induction on $n$. When $n = 0$ both sides evaluate to 1. Assume the claim holds for $n = m$ and consider the case $n = m + 1$.

$$
\begin{aligned}
(x+y)^{m+1} &= (x+y)\cdot(x+y)^m && (10.1)\\
&= (x+y)\cdot\sum_{k=0}^{m}\binom{m}{k}x^k y^{m-k} && (10.2)\\
&= x\cdot\sum_{k=0}^{m}\binom{m}{k}x^k y^{m-k} + y\cdot\sum_{k=0}^{m}\binom{m}{k}x^k y^{m-k} && (10.3)\\
&= \sum_{k=0}^{m}\binom{m}{k}x^{k+1} y^{m-k} + \sum_{k=0}^{m}\binom{m}{k}x^k y^{m+1-k} && (10.4)\\
&= \sum_{k=1}^{m+1}\binom{m}{k-1}x^{k} y^{m+1-k} + \sum_{k=0}^{m}\binom{m}{k}x^k y^{m+1-k} && (10.5)\\
&= \left(x^{m+1}+\sum_{k=1}^{m}\binom{m}{k-1}x^{k} y^{m+1-k}\right) + \left(y^{m+1}+\sum_{k=1}^{m}\binom{m}{k}x^k y^{m+1-k}\right) && (10.6)\\
&= x^{m+1}+y^{m+1}+\sum_{k=1}^{m}\left(\binom{m}{k-1}+\binom{m}{k}\right)x^k y^{m+1-k} && (10.7)\\
&= x^{m+1}+y^{m+1}+\sum_{k=1}^{m}\binom{m+1}{k}x^k y^{m+1-k} && (10.8)\\
&= \sum_{k=0}^{m+1}\binom{m+1}{k}x^k y^{m+1-k}. && (10.9)
\end{aligned}
$$

Here (5) follows from (4) by noting that

$$\sum_{k=0}^{m} f(k) = \sum_{k=1}^{m+1} f(k-1)$$

and (8) follows from (7) by Pascal's rule. The other steps are simple algebraic manipulation. This completes the proof by induction. $\square$

The binomial theorem can be used to immediately derive an identity we have seen before: By substituting $x = y = 1$ into the theorem we get

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

Here is another interesting calculation: Putting $x = -1$ and $y = 1$ yields

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

This implies

$$\sum_{k \text{ odd}} \binom{n}{k} = \sum_{k \text{ even}} \binom{n}{k} = 2^{n-1}.$$

This means that the number of odd-size subsets of an $n$-element set $A$ is the same as the number of even-size subsets, and equals $2^{n-1}$. This can be proved by a combinatorial argument as follows: Fix an element $a \in A$ and note that the number of subsets of $A \setminus \{a\}$ is $2^{n-1}$. There is a bijective map between subsets of $A \setminus \{a\}$ and odd-size subsets of $A$, as follows: Map an odd-sized subset of $A \setminus \{a\}$ to itself, and map an even-sized subset $B \subseteq A \setminus \{a\}$ to $B \cup \{a\}$. Observe that this is a bijection and conclude that the number of odd-sized subsets of $A$ is $2^{n-1}$. Even-size subsets can be treated similarly, or by noting that their number is $2^n$ minus the number of odd-size ones.

# Chapter 11

# The Inclusion-Exclusion Principle

## 11.1 Statement and proof of the principle

We have seen the sum principle that states that for $n$ pairwise disjoint sets $A_1, A_2, \dots, A_n$,

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i|.$$

What happens when the sets are not pairwise disjoint? We can still say something. Namely, the sum $\sum_{i=1}^{n} |A_i|$ counts every element of $\bigcup_{i=1}^{n} A_i$ at least once, and thus even with no information about the sets we can still assert that

$$\left|\bigcup_{i=1}^{n} A_i\right| \leq \sum_{i=1}^{n} |A_i|.$$

However, with more information we can do better. For a concrete example, consider a group of people, 10 of whom speak English, 8 speak French, and 6 speak both languages. How many people are in the group? We can sum the number of English- and French-speakers, getting $10 + 8 = 18$. Clearly, the bilinguals were counted twice, so we need to subtract their number, getting the final answer $18 - 6 = 12$. This argument can be carried out essentially verbatim in a completely general setting, yielding the following formula:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

What if there are three sets? Suppose in addition to the above English and French speakers, we have 14 German-language enthusiasts, among which 8 also speak English, 5 speak French, and 2 speak all three languages. How many people are there now? We can reason as follows: The sum $10 + 8 + 14 = 32$ counts the people speaking two languages twice, so we should subtract their number, getting $32 - 6 - 8 - 5 = 13$. But now the trilinguals have not been counted: They were counted three times in the first sum, and then subtracted three times as part of the bilinguals. So the final answer is obtained by adding their number: $13 + 2 = 15$. In general,

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

In the case of arbitrarily many sets we obtain the inclusion-exclusion principle:

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n}(-1)^{k-1} \sum_{1\le i_1<i_2<\cdots<i_k\le n} |A_{i_1} \cap A_{i_2} \cap \ldots A_{i_k}| .$$

*Proof.* Each element in $\bigcup_{i=1}^{n} A_i$ is counted exactly once on the left side of the formula. Consider such an element $a$ and let the number of sets $A_i$ that contain $a$ be $j$. Then $a$ is counted

$$\binom{j}{1} - \binom{j}{2} + \ldots + (-1)^{j-1}\binom{j}{j}$$

times on the right side. But recall from our exploration of binomial coefficients that

$$\sum_{i=0}^{j}(-1)^i\binom{j}{i} = \sum_{i=0}^{j}(-1)^{i-1}\binom{j}{i} = -1 + \sum_{i=1}^{j}(-1)^{i-1}\binom{j}{i} = 0,$$

which implies

$$\binom{j}{1} - \binom{j}{2} + \ldots + (-1)^{j-1}\binom{j}{j} = 1,$$

meaning that $a$ is counted exactly once on the right side as well. This establishes the inclusion-exclusion principle. □

## 11.2 Derangements

Given a set $A = \{a_1, a_2 \ldots, a_n\}$, we know that the number of bijections from $A$ to itself is $n!$. How many such bijections are there that map no element $a \in A$ to itself? That is, how many bijections are there of the form $f : A \to A$, such that $f(a) \neq a$ for all $a \in A$. These are called *derangements*, or bijections with no *fixed points*.

We can reason as follows: Let $S_i$ be the set of bijections that map the $i$-th element of $A$ to itself. We are the looking for the quantity

$$n! - \left|\bigcup_{i=1}^{n} S_i\right| .$$

By the inclusion-exclusion principle, this is

$$n! - \sum_{k=1}^{n}(-1)^{k-1} \sum_{1\le i_1<i_2<\cdots<i_k\le n} |S_{i_1} \cap S_{i_2} \cap \ldots \cap S_{i_k}| .$$

Consider an intersection $S_{i_1} \cap S_{i_2} \cap \ldots \cap S_{i_k}$. Its elements are the permutations that map $a_{i_1}, a_{i_2}, \ldots, a_{i_k}$ to themselves. The number of such permutations is $(n-k)!$, hence $|S_{i_1} \cap S_{i_2} \cap \ldots \cap S_{i_k}| = (n-k)!$. This allows expressing the number of derangements as

$$
\begin{aligned}
n! - \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} (n-k)! &= n! - \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n-k)! \\
&= \sum_{k=0}^{n} (-1)^{k} \binom{n}{k} (n-k)! \\
&= \sum_{k=0}^{n} (-1)^{k} \frac{n!}{k!} \\
&= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.
\end{aligned}
$$

Now, $\sum_{k=0}^{n} \frac{(-1)^k}{k!}$ is the beginning of the Maclaurin series of $e^{-1}$. (No, you are not required to know this for the exam.) This means that as $n$ gets larger, the number of derangements rapidly approaches $n!/e$. In particular, if we just pick a random permutation of a large set, the chance that it will have no fixed points is about $1/e$. Quite remarkable, isn't it!?

# Chapter 12

# The Pigeonhole Principle

## 12.1 Statements of the principle

In we put more than $n$ pigeons into $n$ pigeonholes, at least one pigeonhole will house two or more pigeons. This trivial observation is the basis of ingenious combinatorial arguments, and is the subject of this chapter. Let's begin with the various guises of the pigeonhole principle that are encountered in combinatorics.

**Basic form.** If $m$ objects are put in $n$ boxes and $n < m$, then at least one box contains at least two objects. The one-line proof is by contradiction: If every box contains at most one object, there are at most $n \cdot 1 = n$ objects. A more rigorous formulation of the principle is as follows: Given two sets $A$ and $B$, with $|A| = m > n = |B|$, for any function $f : A \to B$ there exists $b \in B$ such that

$$\big|\{x \in A : f(x) = b\}\big| > 1.$$

**General form.** If $m$ objects are put in $n$ boxes, then at least one box contains at least $\lceil m/n \rceil$ objects. The proof is again by contradiction: If every box contains at most $\lceil m/n \rceil - 1 < m/n$ objects, there are less than $n(m/n) = m$ objects. The more rigorous formulation is: Given two sets $A$ and $B$, for any function $f : A \to B$ there exists $b \in B$ such that

$$\big|\{x \in A : f(x) = b\}\big| \geq \left\lceil \frac{m}{n} \right\rceil.$$

**Dijkstra's form.** For a nonempty finite collection of integers (not necessarily distinct), the maximum value is at least the average value. It is a good exercise to verify that this is equivalent to the general form above.

## 12.2 Simple applications

Let's begin with some easy applications of the pigeonhole principle.

**First application.** There are two San Franciscans with the exact same number of hairs on their heads. Indeed, according to P&G Hair Facts, the average person's head has about 100,000 hairs, while "some people have as many as 150,000." So it seems safe to bet that every San Franciscan has at most 700,000 hairs on his or her head. On the other hand, the year 2000 US Census counted 776,733 San Francisco residents. The pigeonhole principle implies that at least two of them have the exact same number of hairs.

**Second application.** At a cocktail party with six or more people, there are three mutual acquaintances or three mutual strangers. Indeed, pick an arbitrary person $a$. By the pigeonhole principle, among the other five or more people, either there are three of $a$'s acquaintances, or three people who are strangers to $a$. Let's say there are three that are $a$'s acquaintances, the other case is analogous. If those three are mutual strangers we are done. Otherwise there are two among them, call them $b$ and $c$, who know each other. Then $a$, $b$ and $c$ are mutual acquaintances and we are done.

**Third application.** Consider an infinite two-dimensional plane, every point of which is colored either red or blue; then there are two points one yard apart that are the same color. Indeed, take an arbitrary equilateral triangle with a side length of one yard. By the pigeonhole principle, two of its vertices have the same color.

**Fourth application.** Consider the numbers $1, 2, \ldots, 2n$, and take any $n+1$ of them; then there are two numbers in this sample that are coprime. Indeed, consider the pairs $\{1, 2\}, \{3, 4\}, \ldots, \{2n-1, 2n\}$. By the pigeonhole principle, both numbers from one of these pairs are in the sample. These numbers differ by 1 and are thus coprime. (This follows from the same argument as in Euclid's proof of the infinity of primes.)

## 12.3 Advanced applications

The following lemma comes from a classical 1935 paper by Paul Erdös and George Szekeres titled "A combinatorial problem in geometry":

**Lemma 12.3.1.** *In any ordered sequence of $n^2+1$ distinct real numbers $a_1, a_2, \ldots, a_{n^2+1}$, there is either a monotone increasing subsequence of length $n+1$ or a monotone decreasing subsequence of length $n+1$. Namely, there is a set of indices $1 \le i_1 < i_2 < \cdots < i_{n+1} \le n^2+1$, such that either $a_{i_1} > a_{i_2} > \cdots > a_{i_{n+1}}$ or $a_{i_1} < a_{i_2} < \cdots < a_{i_{n+1}}$.*

*Proof.* For $1 \le i \le n^2+1$, let $\eta_i$ be the length of the longest monotone increasing subsequence that starts at $a_i$. If some $\eta_i > n$, we are done. Otherwise, by the pigeonhole principle, there exists $1 \le j \le n$, and some set $i_1 < i_2 < \cdots < i_m$ of size $m \ge \lceil (n^2+1)/n \rceil = n+1$, such that $\eta_{i_1} = \eta_{i_2} = \cdots = \eta_{i_m} = j$. Now, consider two numbers $a_{i_k}$ and $a_{i_{k+1}}$. If $a_{i_k} < a_{i_{k+1}}$, we get an increasing subsequence starting at $a_{i_k}$ of length $j+1$, which is a contradiction. Hence $a_{i_k} > a_{i_{k+1}}$ in particular, and

$a_{i_1} > a_{i_2} > \cdots > a_{i_m}$ in general, giving us a decreasing subsequence of length at least $n+1$. $\square$

Here is another pigeonhole gem, the last one for today:

**Proposition 12.3.2.** *Given a sequence of $n$ not necessarily distinct integers $a_1, a_2, \ldots, a_n$, there is a nonempty consecutive subsequence $a_i, a_{i+1}, \ldots, a_j$ whose sum $\sum_{m=i}^{j} a_m$ is a multiple of $n$. (The subsequence might consist of a single element.)*

*Proof.* Consider the collection

$$\left(\sum_{i=1}^{0} a_i, \sum_{i=1}^{1} a_i, \sum_{i=1}^{2} a_i, \ldots, \sum_{i=1}^{n} a_i\right).$$

This collection has size $n+1$ and its first element is the empty sum $\sum_{i=1}^{0} a_i = 0$. There are only $n$ possible remainders modulo $n$, thus by the pigeonhole principle, there are two numbers in the above collection of size $n+1$ that leave the same remainder. Let these be $\sum_{i=1}^{l} a_i$ and $\sum_{i=1}^{k} a_i$, with $l < k$. By a lemma we once proved, it follows that

$$n \Big| \left(\sum_{i=1}^{k} a_i - \sum_{i=1}^{l} a_i\right),$$

which implies

$$n \Big| \sum_{i=l+1}^{k} a_i.$$

$\square$