# Counting

The next major topic of the course is probability theory. Suppose you toss a fair coin a thousand times. How likely is it that you get exactly 500 heads? And what about 1000 heads? It turns out that the chances of 500 heads are roughly 2.5%, whereas the chances of 1000 heads are so infinitesimally small that we may as well say that it is impossible. But before you can learn to compute or estimate odds or probabilities you must learn to count! That is the subject of this note.

## 1 Counting Sequences

We start by considering a simple scenario. We pick $k \leq n$ elements from an $n$-element set $S = \{1, 2, \ldots, n\}$ one at a time while removing the sampled element from $S$. This type of sampling is called *sampling without replacement*. We wish to count the number of different ways to do this, taking into account the order in which the elements are picked. For example, when $k = 2$, picking 1 and then 2 is considered a different outcome from picking 2 followed by 1. Another way to ask the question is this: we wish to form an ordered sequence of $k$ distinct elements, where each element is picked from the set $S$. How many different such ordered sequences are there?

If we were dealing cards, the set would be $S = \{1, \ldots, 52\}$, where each number represents a card in a deck of 52 cards. Picking an element of $S$ in this case refers to dealing one card. Note that once a card is dealt, it is no longer in the deck and so it cannot be dealt again. So the hand of $k$ cards that are dealt consists of $k$ distinct elements from the set $S$.

For the first card, it is easy to see that we have 52 distinct choices. But now the available choices for the second card depend upon what card we picked first. The crucial observation is that regardless of which card we picked first, there are exactly 51 choices for the second card. So the total number of ways of choosing the first two cards is $52 \times 51$. Reasoning in the same way, there are exactly 50 choices for the third card, regardless of our choices for the first two cards. It follows that there are exactly $52 \times 51 \times 50$ sequences of three cards. In general, the number of sequences of $k$ cards is $52 \times 51 \times \cdots \times [52 - (k - 1)]$.

This is an example of the First Rule of Counting:

**First Rule of Counting:** If an object can be made by a succession of $k$ choices, where there are $n_1$ ways of making the first choice, and *for every* way of making the first choice there are $n_2$ ways of making the second choice, and *for every* way of making the first and second choice there are $n_3$ ways of making the third choice, and so on up to the $n_k$-th choice, then the total number of distinct objects that can be made in this way is the product $n_1 \times n_2 \times n_3 \times \cdots \times n_k$.

Here is another way of picturing the First Rule of Counting. Consider the following tree:
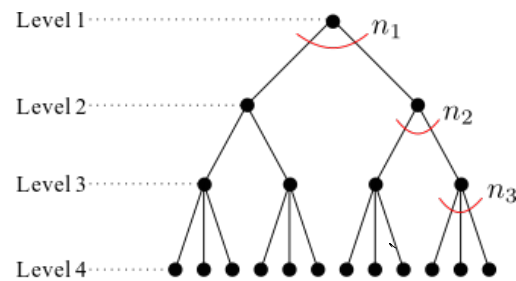
It has branching factor $n_1$ at the root, $n_2$ at every vertex at the second level,..., $n_k$ at every vertex at the $k$-th level. Each path from the root to a vertex at level $k+1$ (i.e., a leaf vertex) represents one possible way of making the object by making a succession of $k$ choices. So the number of distinct objects that can be made is equal to the number of leaves in the tree. Moreover, the number of leaves in the tree is the product $n_1 \times n_2 \times n_3 \times \cdots \times n_k$. For example, if $n_1 = 2$, $n_2 = 2$, and $n_3 = 3$, then there are 12 leaves (i.e., outcomes).

## 2   Counting Sets

Consider a slightly different question. We would like to pick $k$ distinct elements of $S = \{1, 2, \ldots, n\}$ (i.e., without repetition), but we do not care about the order in which we picked the $k$ elements. For example, picking elements $1, \ldots, k$ is considered the same outcome as picking elements $2, \ldots, k$ and picking 1 as the last ($k$-th element). Now how many ways are there to choose these elements to obtain the same outcome?

When dealing a hand of cards, say a poker hand, it is often more natural to count the number of distinct hands (i.e., the set of 5 cards dealt in the hand), rather than the order in which they were dealt. As we have seen in Section 1, if we are considering order, there are $52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 = \frac{52!}{47!}$ outcomes. But how many distinct hands of 5 cards are there? Here is another way of asking the question: each such 5 card hand is just a subset of $S$ of cardinality 5. So we are asking how many 5 element subsets of $S$ are there?

Here is a clever trick for counting the number of distinct subsets of $S$ with exactly 5 elements. Create a bin corresponding to each such 5 element subset. Now take all the sequences of 5 cards and distribute them into these bins in the natural way: each sequence gets placed in the bin corresponding to the set of 5 elements in the sequence. Thus, if the sequence is $(2, 1, 3, 5, 4)$, then it is placed in the bin labeled $\{1, 2, 3, 4, 5\}$. How many sequences are placed in each bin? The answer is exactly 5!, since there are exactly 5! different ways to order 5 cards.

Recall that our goal was to compute the number of 5 element subsets, which now corresponds to the number of bins. We know that there are $\frac{52!}{47!}$ distinct 5-card sequences, and there are 5! sequences placed in each bin. The total number of bins is therefore $\frac{52!}{47!5!}$.

This quantity $\frac{n!}{(n-k)!k!}$ is used so often that there is special notation for it: $\binom{n}{k}$, pronounced "$n$ choose $k$," and it's known as the *binomial coefficient*. This is the number of ways of picking $k$ distinct elements from $S$, where the order of placement does not matter. Equivalently, it's the number of ways of choosing $k$ objects out of a total of $n$ distinct objects, where the order of the choices does not matter.

The trick we used above is actually our Second Rule of Counting:

**Second Rule of Counting:** Assume an object is made by a succession of choices, and the order in which the choices are made does not matter. Let $A$ be the set of ordered objects and let $B$ be the set of unordered objects. If there exists an $m$-to-1 function $f$ from $A$ to $B$, we can count the number of ordered objects (pretending that the order matters) and divide by $m$ (the number of ordered objects per unordered objects) to obtain $|B|$, the number of unordered objects.

Note that we are assuming that the number of ordered objects is the same for every unordered object; the rule cannot be applied otherwise. Here is another way of picturing the Second Rule of Counting:
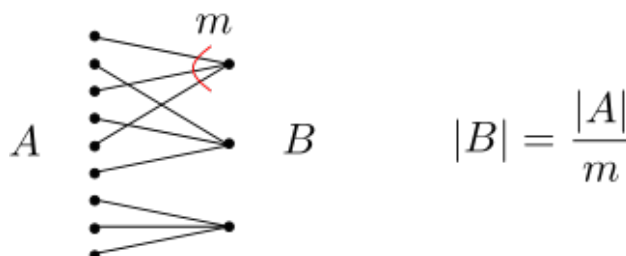


The function $f$ simply places the ordered outcomes into bins corresponding to the unordered outcomes. In our poker hand example, $f$ will map 5! elements in the domain of the function (the set of ordered 5 card outcomes) to one element in the range (the set of 5-element subsets of $S$). The number of elements in the range of the function is therefore $\frac{52!}{47!5!}$.
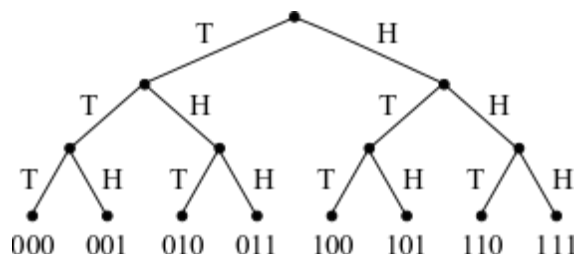
# 3 Sampling with Replacement

Sometimes we wish to consider a different scenario of sampling where after we choose an element from $S = \{1, 2, \ldots, n\}$, we throw it back into $S$ so that we can choose it again. Assume that we are still picking $k$ elements out of $S$ one at a time and that order matters. How many distinct sequences of $k$ elements can we obtain under this new sampling scheme? We can use the First Rule of Counting. Since we have $n$ choices in each trial, $n_1 = n_2 = \cdots = n_k = n$. Then we have a grand total of $n^k$ sequences.

This type of sampling is called *sampling with replacement*; multiple trials can have the same outcome. Card dealing is a type of *sampling without replacement*, since two trials cannot both have the same outcome (one card cannot be dealt twice).

## 3.1 Coin Tosses

Let us return to coin tosses. How many different outcomes are there if we toss a coin $k$ times? By an outcome, we mean a length-$k$ sequence consisting of heads and tails. Suppose we represent each outcome by a binary string of length $k$ where 0 corresponds to a tail ($T$) and 1 a head ($H$). For example, the string 001 would represent an outcome of 2 tails followed by a head. The following tree, in which the leaves are in 1-to-1 correspondence with the set of possible outcomes, illustrates the experiment for $k = 3$:



Here $S = \{0, 1\}$. This is a case of sampling with replacement; multiple coin flips could result in tails (we could pick the element 0 from the set $S$ in multiple trials). Order also matters; e.g., strings 001 and 010 are considered different outcomes. By the reasoning above (using the First Rule of Counting) we have a total of $n^k = 2^k$ distinct outcomes, since $n = 2$.

## 3.2 Rolling Dice

Let's say we roll two dice, so $k = 2$ and $S = \{1,2,3,4,5,6\}$. How many possible outcomes are there? In this setting, ordering matters; obtaining 1 with the first die and 2 with the second is different from obtaining 2 with the first and 1 with the second. We are sampling with replacement, since we can obtain the same result on both dice.

The setting is therefore the same as the coin flipping example above (order matters and we are sampling with replacement), so we can use the First Rule of Counting in the same manner. The number of distinct outcomes is therefore $n^2 = 6^2 = 36$.

## 3.3 Sampling with Replacement, but where Order Does Not Matter

Say you have unlimited quantities of apples, bananas and oranges. You want to select 5 pieces of fruit to make a fruit salad. How many ways are there to do this? In this example, $S = \{1,2,3\}$, where 1 represents apples, 2 represents bananas, and 3 represents oranges. $k = 5$ since we wish to select 5 pieces of fruit. Ordering does not matter; selecting an apple followed by a banana will lead to the same salad as a banana followed by an apple.

This scenario is trickier to analyze. It may seem natural to apply the Second Rule of Counting because order does not matter. Let's consider this method. We first pretend that order matters and observe the number of ordered objects is $3^5$ as discussed above. How many ordered options are there for every unordered option? The problem is that this number differs depending on which unordered object we are considering. Let's say the unordered object is an outcome with 5 bananas. There is only one such ordered outcome. But if we are considering 4 bananas and 1 apple, there are 5 such ordered outcomes (represented as 12222, 21222, 22122, 22212, 22221).

Now that we see the Second Rule of Counting will not help, can we look at this problem in a different way? Let us first generalize back to our original setting: we have a set $S = \{1,2,\ldots,n\}$ and we would like to know how many ways there are to choose multisets (sets with repetition) of size $k$. Remarkably, we can model this problem in terms of binary strings, as described below.

Assume we have one bin for each element of $S$, so $n$ bins in total. For example, if we selected 2 apples and 1 banana, bin 1 would have 2 elements and bin 2 would have 1 element. In order to count the number of multisets, we need to count how many different ways there are to fill these bins with $k$ elements. We don't care about the order of the bins themselves, just how many of the $k$ elements each bin contains. Let's represent each of the $k$ elements by a 0 in the binary string, and separations between bins by a 1. The following picture illustrates an example of placement where $S = \{1,\ldots,5\}$ and $k = 4$:

$$\boxed{OO}\,\boxed{\phantom{O}}\,\boxed{O}\,\boxed{\phantom{O}}\,\boxed{O}$$

00110110

Counting the number of multisets is now equivalent to counting the number of placements of the $k$ 0's. We have just reduced what seemed like a very complex problem to a question about a binary string, simply by looking at it from a different perspective!

↑$k$ 0's  ↱$n-1$ 1's to seperate bins

How many ways can we choose these locations? The length of our binary string is $k + n - 1$, and we are choosing which $k$ locations should contain 0's. The remaining $n - 1$ locations will contain 1's. Once we pick a location for one zero, we cannot pick it again; repetition is not allowed. Picking location 1 followed by location 2 is the same as picking location 2 followed by location 1, so ordering does not matter. It follows

that all we wish to compute is the number of ways of picking $k$ elements from $k+n-1$ elements, without replacement and where the order of placement does not matter. This is given by $\binom{n+k-1}{k}$, as discussed in Section 2. This is therefore the number of ways in which we can choose multisets of size $k$ from set $S$.

Returning to our example above, the number of ways of picking 5 pieces of fruit is exactly $\binom{3+5-1}{5} = \binom{7}{5}$.

Notice that we started with a problem which seemed very different from previous examples, but, by viewing it from a certain perspective, we were able to use previous techniques (those used in counting sets) to find a solution!

In a sense, we have found the zeroth rule of counting:

**Zeroth Rule of Counting:** If a set $A$ can be placed into a one-to-one correspondance with a set $B$ (i.e. you can find a bijection between the two — an invertible pair of maps that map elements of $A$ to elements of $B$ and vice-versa), then $|A| = |B|$.

This is the very heart of what it means to count and is key to many combinatorial arguments as we will explore further in the next section.

## 4  Combinatorial Proofs

Combinatorial proofs are interesting because they rely on intuitive counting arguments rather than tedious algebraic manipulation. They often feel like proofs by stories — the same story, told from multiple points of view.

As a warmup, let's prove the Binomial Theorem using counting arguments:

**Theorem 10.1 (Binomial Theorem).** *For all $n \in \mathbb{N}$,*

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

*Proof.* The left-hand side can be written as the $n$-fold product $(a+b)(a+b)\cdots(a+b)$. Label these $n$ factors as $f_1, f_2, \ldots, f_n$. When $(a+b)(a+b)\cdots(a+b)$ is expanded out via multiplication, it will result in a sum over $2^n$ monomials, and to each monomial in the sum, each factor $f_i$ contributes either an $a$ or a $b$. For example, in $(a+b)(a+b) = a^2 + ab + ba + b^2$, $a^2$ is formed by multiplying $a$ from $f_1$ and $a$ from $f_2$; $ab$ is formed by multiplying $a$ from $f_1$ and $b$ from $f_2$; $ba$ is formed by multiplying $b$ from $f_1$ and $a$ from $f_2$; $b^2$ is formed by multiplying $b$ from $f_1$ and $b$ from $f_2$. In general, the monomial $a^k b^{n-k}$ is obtained when $k$ copies of $a$'s and $n-k$ copies of $b$'s are multiplied. There are exactly $\binom{n}{k}$ ways to choose $k$ copies of $a$'s and $n-k$ copies of $b$'s from the $n$ factors $f_1, f_2, \ldots, f_n$. $\qquad \square$

This well-known result has numerous applications. For example, the following combinatorial identity can be easily derived using Theorem 10.1:

**Corollary 10.1.** *For all $n \in \mathbb{N}$,*

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

*Proof.* Set $a = -1$ and $b = 1$ in Theorem 10.1. $\qquad \square$

In what follows, we will consider more sophisticated versions of combinatorial proofs. The key idea is to show that the two sides of a given equation correspond to two different ways to count the same kind of objects. In more advanced settings, the goal is to establish a bijection between the set of objects counted on the left-hand side of an identity with the set of objects counted on the right-hand side. We illustrate these concepts below.

## 4.1 Combinatorial Identities

Using combinatorial arguments, we can prove complex facts, such as

$$\binom{n}{k+1} = \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k}{k}, \qquad \binom{n}{k+1} = \sum_{i=k}^{n-1} \binom{i}{k} \tag{1}$$

which is known as the *hockey-stick identity*. You can directly verify this identity by algebraic manipulation. But you can also do this by interpreting what each side means as a combinatorial process. The left-hand side of (1) is just the number of ways of choosing a subset $X$ with $k+1$ elements from a set $S = \{1,\ldots,n\}$. Let us think about a different process that results in the choice of a subset with $k+1$ elements. We start by picking the lowest-numbered element of $X$. If we picked 1, to finish constructing $X$, we must now choose $k$ elements out of the remaining $n-1$ elements of $S$ greater than 1; this can be done in $\binom{n-1}{k}$ ways. If instead the lowest-numbered element we picked is 2, then we have to choose $k$ elements from the remaining $n-2$ elements of $S$ greater than 2, which can be done in $\binom{n-2}{k}$ ways. Moreover all these subsets are distinct from those where the lowest-numbered element was 1. So we should add the number of ways of choosing each to the grand total. Proceeding in this way, we split up the process into cases according to the first (i.e., lowest-numbered) object we select, to obtain:

$$
\textbf{First element selected is} \begin{cases}
\text{element } 1, & \text{leading to } \binom{n-1}{k} \text{ remaining ways,} \\
\text{element } 2, & \text{leading to } \binom{n-2}{k} \text{ remaining ways,} \\
\text{element } 3, & \text{leading to } \binom{n-3}{k} \text{ remaining ways,} \\
\quad \vdots & \qquad \vdots \\
\text{element } (n-k), & \text{leading to } \binom{k}{k} \text{ remaining ways.}
\end{cases}
$$

These factors correspond to the terms on the right-hand side of (1). (Note that the lowest-numbered object we select cannot be higher than $n-k$ as we have to select $k+1$ distinct objects.)

Another combinatorial identity we will prove is

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n. \tag{2}$$

This identity actually follows immediately from setting $a = 1$ and $b = 1$ in the Binomial Theorem above, but we will provide a more direct combinatorial proof. Suppose we have a set $S$ with $n$ distinct elements. On the left-hand side of (2), $\binom{n}{i}$ counts the number of ways of choosing a subset of $S$ of size exactly $i$; so the sum on the left-hand side counts the total number of subsets (of any size) of $S$.

We claim that the right-hand side of (2) does indeed also count the total number of subsets. To see this, just identify a subset with an $n$-bit vector, where in each position $j$ we put a 1 if the $j$-th element is in the subset, and a 0 otherwise. So the number of subsets is equal to the number of $n$-bit vectors, which is $2^n$ (there are 2 options for each bit).

Let us look at an example, where $S = \{1,2,3\}$ (so $n = 3$). Enumerate all $2^3 = 8$ possible subsets of $S$: $\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}$. The eight 3-bit vectors corresponding to these subsets are

Table 1: Permutations of $\{1, 2, 3\}$ and the number of fixed points.

| $(\pi_1, \pi_2, \pi_3)$ | # fixed points |
|---|---|
| $(1,2,3)$ | 3 |
| $(1,3,2)$ | 1 |
| $(2,1,3)$ | 1 |
| $(2,3,1)$ | 0 |
| $(3,1,2)$ | 0 |
| $(3,2,1)$ | 1 |

$(0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), (1,1,1)$, respectively. On the left-hand side of equation (2), the term $\binom{3}{0}$ counts the number of ways to choose a subset of $S$ with 0 elements; there is only one such subset, namely the empty set $\varnothing = \{\}$. There are $\binom{3}{1} = 3$ ways of choosing a subset with 1 element, $\binom{3}{2} = 3$ ways of choosing a subset with 2 elements, and $\binom{3}{3} = 1$ way of choosing a subset with 3 elements (namely, the subset consisting of every element of $S$). Summing, we get $1 + 3 + 3 + 1 = 8$, as expected.

*↷ No fix point perms*

## 4.2 Permutations and Derangements

**Question:** Suppose we collect the homeworks of $n$ students, randomly shuffle them, and return them to the students. If we repeated this experiment many times, then on average how many students would receive their own homework?

At first this may seem like a rather challenging problem, but you will be able to answer this question in a few lines later in the course once you have some basic tools of probability theory under your belt. For now, let's translate the question into a mathematical statement and consider a related counting problem.

Label the homeworks as $1, 2, \ldots, n$ and let $\pi_i$ denote the homework that is returned to the $i$-th student. Then, the vector $\pi = (\pi_1, \ldots, \pi_n)$ corresponds to a permutation of the set $\{1, \ldots, n\}$. Note that $\pi_1, \ldots, \pi_n \in \{1, 2, \ldots, n\}$ are all distinct, so each element in $\{1, \ldots, n\}$ appears exactly once in the permutation $\pi$. In total, there are $n!$ distinct permutations of $\{1, \ldots, n\}$.

In this setting, the $i$-th student receives their own homework if and only if $\pi_i = i$. Then the question "How many students receive their own homework?" translates into the question of how many indices ($i$'s) satisfy $\pi_i = i$. These are known as fixed points of the permutation $\pi$. To illustrate the idea concretely, let us consider the example $n = 3$. Table 1 shows a complete list of all $3! = 6$ permutations of $\{1, 2, 3\}$, together with the corresponding number of fixed points.

We now consider an interesting counting problem concerning derangements, which are defined as follows.

**Definition 10.1** (Derangement). *A permutation with no fixed points is called a* derangement.

Let $D_n$ denote the number of derangements of $\{1, 2, \ldots, n\}$. As can be seen in Table 1, $D_3 = 2$; specifically, the derangements are $(2, 3, 1)$ and $(3, 1, 2)$. Our goal is to find a formula for $D_n$ as a function of $n$ for all $n \in \mathbb{Z}^+$. We will learn two different ways to solve this problem.

The first method involves a combinatorial proof. For concreteness, let's consider the case of $n = 4$. Table 2a lists all derangements of $\{1, 2, 3, 4\}$, with each row corresponding to a derangement. In the first derangement, note that the entries shown in blue corresponds to a derangement of $\{1, 2\}$. In the second and third derangements, the entries shown in red are in 1-to-1 correspondence with the derangements of $\{1, 2, 3\}$, shown in Table 2b. Note that the red 4 in Table 2a satisfies the same constraint as the number 3 in Table 2b; specifically, $\pi_3 \neq 4$ in the former and $\pi_3 \neq 3$ in the latter. In summary, the above discussion suggests that $D_4$ is related to $D_2$ and $D_3$. The following theorem formalizes this relationship.

Table 2: List of derangements. (a) $n = 4$. (b) $n = 3$.

(a)

| $\pi_1$ | $\pi_2$ | $\pi_3$ | $\pi_4$ |
|---|---|---|---|
| **2** | **1** | 4 | 3 |
| **2** | **4** | **1** | 3 |
| **4** | **1** | **2** | 3 |
| 4 | 3 | 2 | 1 |
| 3 | 4 | 2 | 1 |
| 2 | 3 | 4 | 1 |
| 3 | 4 | 1 | 2 |
| 4 | 3 | 1 | 2 |
| 3 | 1 | 4 | 2 |

(b)

| $\pi_1$ | $\pi_2$ | $\pi_3$ |
|---|---|---|
| 2 | 3 | 1 |
| 3 | 1 | 2 |

**Theorem 10.2.** *For an arbitrary positive integer $n \geq 3$, the number $D_n$ of derangements of $\{1,\ldots,n\}$ satisfies*

$$D_n = (n-1)(D_{n-1} + D_{n-2}). \tag{3}$$

*Proof.* We provide a combinatorial proof by establishing relevant bijections. In a derangement $(\pi_1,\ldots,\pi_n)$ of $\{1,\ldots,n\}$, suppose $\pi_n = j \in \{1,\ldots,n-1\}$; i.e., $j$ is what $n$ maps to under the permutation. There are $n-1$ choices for $j$, and this gives rise to the multiplicative factor $(n-1)$ in (3). (The case of $j = n$ is excluded, since a derangement cannot have $\pi_n = n$.) For a given value of $j \in \{1,\ldots,n-1\}$, there are two cases to consider:

Case 1: If $\pi_j = n$, then $j$ and $n$ get swapped, and there is a 1-to-1 correspondence between the set of derangements of the remaining elements and the set of derangements of $\{1,\ldots,n-2\}$. This gives rise to $D_{n-2}$.

Case 2: If $\pi_j \neq n$, then $n$ satisfies the same constraint as $j$, so there is a 1-to-1 correspondence between the set of rearrangements of $\{1,\ldots,j-1,j+1,\ldots,n\}$ and the set of derangements of $\{1,\ldots,n-1\}$. This gives rise to $D_{n-1}$.

This completes our derivation of (3). $\square$

Together with the boundary conditions $D_1 = 0$ and $D_2 = 1$ (check this), $D_n$ can be determined efficiently (in linear time in $n$) using the recursion in (3) and memoization. In fact, the recursion can be solved analytically to yield (see Exercise 2) an explicit formula:

$$D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}. \tag{4}$$

Coming up with the recursion in (3) and solving it require some creativity. As mentioned earlier, there is an alternate way to tackle this counting problem, which turns out to be a more straightforward approach. This alternate method utilizes a powerful tool called the principle of inclusion-exclusion, which we now describe.

# 5  The Principle of Inclusion-Exclusion

In this section we will learn a useful formula that can be employed to count objects without over-counting. We start with a familiar example. Let $A_1$ and $A_2$ be two subsets of the same finite set $A$, and suppose we want

to count the number of elements in $A_1 \cup A_2$. If $A_1$ and $A_2$ are disjoint, then clearly $|A_1 \cup A_2| = |A_1| + |A_2|$. If $A_1$ and $A_2$ have common elements, however, then those elements would be counted twice in the previous formula. To correct for this overcounting, we need to subtract $|A_1 \cap A_2|$ from $|A_1| + |A_2|$, thus yielding

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|.$$

The principle of inclusion-exclusion generalizes this concept to an arbitrary number of subsets.

**Theorem 10.3** (Inclusion-Exclusion). *Let $A_1, \ldots, A_n$ be arbitrary subsets of the same finite set A. Then,*

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq \{1,\ldots,n\}:|S|=k} |\cap_{i \in S} A_i|. \tag{5}$$

*All subsets of index i of length k.*

*Remarks:*

1. The inner summation in (5) is over all size-$k$ subsets of $\{1, 2, \ldots, n\}$. More explicitly, (5) can be written as

$$|A_1 \cup \cdots \cup A_n| = \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \sum_{i<j<k<l} |A_i \cap A_j \cap A_k \cap A_l| + \cdots$$
$$+ (-1)^{n-1} |A_1 \cap A_2 \cap \cdots \cap A_n|, \tag{6}$$

   where $\sum_{i<j}$ denotes summing over all $i, j \in \{1, \ldots, n\}$ such that $i < j$, and so on.

2. The standard way to prove Theorem 10.3 is via induction on $n$. Here, we will provide a combinatorial proof.

*Proof.* First, note that if $a \notin A_1 \cup \cdots \cup A_n$, then $a \notin A_i$ for all $i \in \{1, \ldots, n\}$ and hence $a \notin \cap_{i \in S} A_i$ for any $S \subseteq \{1, \ldots, n\}$, so it is not counted by the right-hand side (RHS) of (5). We now show that if $a \in A_1 \cup \cdots \cup A_n$, then it is counted exactly once by the RHS. Define the index set $M \subseteq \{1, \ldots, n\}$ as

$$M = \{i \in \{1, \ldots, n\} \mid a \in A_i\};$$

that is, $a \in A_i$ if and only if $i \in M$. Note that $m := |M| \geq 1$, since $a \in A_1 \cup \cdots \cup A_n$. Clearly, if $S \nsubseteq M$, then $a \notin \cap_{i \in S} A_i$ (i.e., if $S$ contains an index $i$ not in $M$, then $a \notin A_i$, so $a \notin \cap_{i \in S} A_i$). Furthermore, $a \in \cap_{i \in S} A_i$ for any non-empty subset $S \subseteq M$. Therefore, the number of times that $a$ is counted on the RHS of (5) is

$$\sum_{k=1}^{m} (-1)^{k-1} \sum_{S \subseteq M:|S|=k} 1 = \sum_{k=1}^{m} (-1)^{k-1} \binom{m}{k} = 1,$$

where the first equality follows from the fact that the number of distinct $k$-element subsets of $M$ is $\binom{m}{k}$, while the second equality follows from Corollary 10.1. $\square$

## Application: Derangements Revisited

We now illustrate the power of Theorem 10.3 by applying it to derive the formula shown in (4) for the number $D_n$ of derangements of $\{1, \ldots, n\}$.

Let $A_i$ denote the set of all permutations of $\{1, \ldots, n\}$ with $i$ being a fixed point. Then, since there are $n!$ distinct permutations of $\{1, \ldots, n\}$ and $|A_1 \cup A_2 \cup \cdots \cup A_n|$ counts the number of permutations with at least one fixed point, we have

$$D_n = n! - |A_1 \cup A_2 \cup \cdots \cup A_n|. \tag{7}$$

What is the cardinality of $A_i$? We know that $i$ maps to $i$, while the remaining $n-1$ elements can be arranged in an arbitrary way. So, $|A_i| = (n-1)!$. Similarly, for $i \neq j$, $i$ maps to $i$ and $j$ maps to $j$, while the remaining $n-2$ elements can be arranged in an arbitrary way, so we conclude $|A_i \cap A_j| = (n-2)!$. In general, for any subset $S \subseteq \{1, \ldots, n\}$ of size $|S| = k$, $|\cap_{i \in S} A_i| = (n-k)!$, while there are $\binom{n}{k}$ such subsets. Thus, the inclusion-exclusion formula (5) gives

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k}(n-k)! = \sum_{k=1}^{n} (-1)^{k-1} \frac{n!}{k!(n-k)!}(n-k)! = n! \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}.$$

Substituting this result into (7) gives

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

## 6  Exercises

1. Provide a direct combinatorial proof of Corollary 10.1.

2. Using (3) and the boundary conditions $D_1 = 0$ and $D_2 = 1$, derive (4). [Hint: For $n \geq 2$, define $Q_n = D_n - nD_{n-1}$ and obtain a recursion for $Q_n$.]

3. Prove Theorem 10.3 using induction.