1 The Inclusion-Exclusion Theorem

We have already seen the equality $|A \cup B| = |A| + |B| - |A \cap B|$ and its proof by drawing the picture and arguing that the elements in $A \cap B$ get counted twice, and therefore we must subtract $|A \cap B|$.

This naturally extends to 3 sets:

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

We can again argue combinatorially, but we can also give a proof by doing a single inductive step: $|A \cup B \cup C| = |(A \cup B) \cup C| = |A \cup B| + |C| - |(A \cup B) \cap C|$. We now use the formula for $|A \cup B|$ and get

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |(A \cup B) \cap C|$$

Since $|(A \cup B) \cap C| = (A \cap C) \cup (B \cap C)$ by DeMorgan’s Laws, we use again the formula for cardinality of the union applied to the sets $(A \cap C)$ and $(B \cap C)$, getting

$$|(A \cap C) \cup (B \cap C)| = |A \cap C| + |B \cap C| - |(A \cap C) \cap (B \cap C)|$$

Clearly, the last term is simply $|A \cap B \cap C|$. Putting everything together we get the desired equality.

We can extend the equality to the union of $n$ sets:

**Theorem 1** (Inclusion-Exclusion). Let $A_i, i = 1, \cdots, n$ be sets. Then

$$|\bigcup_{i=1}^{n} A_i| = \sum_{1 \leq i \leq n} |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \cdots (-1)^n |\bigcap_{i=1}^{n} A_i|$$
Proof. One can do a proof by induction, along the lines of the arguments we presented to prove the case of \( n = 3 \).

There is also an elegant combinatorial proof: consider an element \( a \) that is a member of exactly \( n \) subsets, and look at the formula above. Each set will contribute 1 to the sum—each \( A_i \) such that \( a \in A_i \) contributes 1, each pair \( \{A_i, A_j\} \) such that \( a \in A_i \) and \( a \in A_j \) will contribute \(-1\) when considering the term with \( A_i \cap A_j \) and so on. The total contribution of all terms can be computed by noting the there will be exactly \( \binom{n}{k} \) \( k \)-tuples of sets that contain \( a \), so for each such \( a \) we will have a total contribution of \( \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \).

A slick way to compute this sum is to use the identity (derived from the Binomial Theorem)

\[
0 = (1 - 1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k 1^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k
\]

We note that the first term of the sum is \( \binom{n}{0} = 1 \), and the remaining of the sum is our expression multiplied by \(-1\). So the expression evaluates to 1. Since this is true for every \( a \), the theorem follows.

Example 2 (Sieve of Eratosthenes). An ancient method to find the prime numbers up to \( n \) is attributed to Eratosthenes of Cyrene, a Greek mathematician, and head of the Library of Alexandria (3rd century BC). You write down the numbers from 1 to \( n \), the cross out every second (starting the count at 1), the cross out every 3rd, and so on. The uncrossed numbers are the primes and 1.

It is clear that one only needs to perform the algorithm till \( \lceil \sqrt{n} \rceil \) divisors (if \( n \) is composite, it has a factor no larger than that) and that we only have to perform the crossing for numbers that are not multiples of previous numbers.

If we only want to count the number of primes, we just do truncating divisions and the Inclusion-Exclusion Theorem.

Let us count the number of primes less or equal to 100:

Let \( A_2 \) be the set of multiples of 2 (excluding 2 itself) less than 100.

\[ |A_2| = \lfloor 100/2 \rfloor = 50 \]

Similarly, let \( A_3 \) be the set of multiples of 3 (excluding 3 itself) less than 100.

\[ |A_3| = \lfloor 100/3 \rfloor = 33 \]

We do similar computations for the multiples of 5 and 7 (note that \( 11^2 > 100 \)).
We then have to compute the cardinalities of the sets of pairs – for example, \( A_2 \cap A_5 \) is the set of numbers that are multiples of both 2, and 5 – these are the multiples of 10, and there are \( \lceil 100/10 \rceil - 10 \) of them. We subtract the cardinalities of all such pairs, then add the cardinality of triples with product less than 100, etc.

It is a messy but straightforward calculation...

1.1 Derangements

Definition 3. An \( n \)-permutation \( \pi() \) can be viewed as a 1-1 function. \( \{1, \cdots n\} \rightarrow \{1, \cdots n\} \). It fixes \( i \) if \( \pi(i) = i \). It is a derangement if no element is fixed - i.e. \( \forall i \ \pi(i) \neq i \).

We will use inclusion-exclusion to compute \( D_n \), the number of derangements of \( n \) objects.

Definition 4 (Another Hatcheck Problem). Consider the hatcheck problem we studied before. It is clearly equivalent to the following: given a set of \( n \) people who checked their hat, they are handed the result of a random permutation of the hats – the \( i \)-th person receives the hat of person \( \pi(i) \), with \( \pi \) chosen from the uniform distribution over permutations.

In our previous problem we showed that the expected number of people who get their own hat back is 1.

The question we ask now is: What is the probability that no person gets their hat back?

It should be clear that the answer is \( D_n/n! \). In order to get a number out of this formula, we have to compute \( D_n \).

Theorem 5 (The number of \( n \)-derangements).

\[
D_n = n! \left[ 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots (-1)^n \frac{1}{n!} \right]
\]

Proof. We will use inclusion-exclusion. Let \( F(i) \) be the set of permutations that fix \( i \), \( F(i,j) \) the set of permutations that fix \( i \) and \( j \), and so on. For example \( F(3,5,17,21) \) is the set of permutations that fix the elements 3, 5, 17, and 21.

It should be clear that \( |F(i)| = (n-1)! \), and, in general, \( |F(i_1,i_2,\cdots i_k)| = (n-k)! \). Moreover, there are exactly \( \binom{n}{k} \) distinct \( k \)-subsets of \( n \), and permutations that fix distinct subsets must be distinct. We
can compute the number of permutations that fix some subset of elements using inclusion-exclusion. The sets whose union we want to compute are the $F(i)$, and the intersections are given by $F$s with several arguments: $F(i, j) = F(i) \cap F(j)$, $F(i, j, k) = F(i) \cap F(j) \cap F(k)$, etc.

So $\text{FIX}$, the number of permutations that fix some element is

$$\text{FIX} = | \bigcup_i F(i) | = \sum_{1 \leq i \leq n} |F(i)| - \sum_{1 \leq i < j \leq n} |F(i, j)| + \sum_{1 \leq i < j < k \leq n} |F(i, j, k)| \cdots (-1)^{n+1} \left( \frac{n}{n} \right) |F(1, 2, \ldots n)|$$

and the number of derangements is

$$n! - \text{FIX} = n! - \sum_{1 \leq i \leq n} \left( \frac{n}{1} \right) |F(i)| + \sum_{1 \leq i < j \leq n} \left( \frac{n}{2} \right) |F(i, j)| \cdots (-1)^n \left( \frac{n}{n} \right) |F(1, 2, \ldots n)|$$

Substituting our formulas for the $F$s, we have

$$n! - \text{FIX} = n! - \left( \frac{n}{1} \right) (n - 1)! + \left( \frac{n}{2} \right) (n - 2)! \cdots (-1)^n \left( \frac{n}{n} \right) 1!$$

Now observe that $\left( \frac{n}{k} \right) (n - k)! = \frac{n!}{k!(n-k)!}$, so, by factoring $n!$ we get

$$D_n = n! \left[ 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} \cdots (-1)^n \frac{1}{n!} \right]$$

Back to the hatcheck problem: the probability that nobody gets their own hat back is $D_n/n!$, which is

$$1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} \cdots (-1)^n \frac{1}{n!}$$

We can get a good estimate for this quantity, by remembering two facts from Calculus:

1. $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$
2. In an alternating series (\( \forall i \, |a_{i+1}| < |a_i| \), and the signs of the terms alternate) the error committed by truncating the series with the \( i \)-th term is bounded by \( |a_{i+1}| \)

We see that the formula for \( D_n/n! \) consists of the first \( n \) terms of the series for \( e-1 \), so its is within \( 1/(n!) \) of \( e-1 \).