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 \( A_1, A_2, \ldots, A_n \),

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

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:

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

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. \( \square \)

### 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

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

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

\[
n! - \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} (n - k)! = n! - \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n - k)!
\]

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

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

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

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?!