8.5 Inclusion-Exclusion

In Section 2.2 we saw that the number of elements in the union of the two sets A and B is the sum of the numbers of elements in the sets minus the number of elements in their intersection. That is,

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

Example 1. *How many positive integers not exceeding* 1000 *are divisible by* 7 *or* 11*?*

Solution. Let A be the set of positive integers not exceeding 1000 that are divisible by 7, and let B be the set of positive integers not exceeding 1000 that are divisible by 11. Then $A \cup B$ is the set of integers not exceeding 1000 that are divisible by either 7 or 11, and $A \cap B$ is the set of integers not exceeding 1000 that are divisible by both 7 and 11. From Section 4.1, we know that among the positive integers not exceeding 1000 there are $\lfloor \frac{1000}{7} \rfloor$ integers divisible by 7 and $\lfloor \frac{1000}{11} \rfloor$ divisible by 11. Because 7 and 11 are relatively prime, the integers divisible by both 7 and 11 are those divisible by $7 \cdot 11$. Consequently, there are $\lfloor \frac{1000}{7 \cdot 11} \rfloor$ positive integers not exceeding 1000 that are divisible by both 7 and 11. It follows that there are

$$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ &= \left\lfloor \frac{1000}{7} \right\rfloor + \left\lfloor \frac{1000}{11} \right\rfloor - \left\lfloor \frac{1000}{7 \cdot 11} \right\rfloor \\ &= 142 + 90 - 12 \\ &= 220 \end{aligned}$$

positive integers not exceeding 1000 that are divisible by either 7 or 11. □

What if there are three sets?

We will now begin our development of a formula for the number of elements in the union of a finite number of sets. The formula we will develop is called *the principle of inclusion-exclusion.* For concreteness, before we consider unions of n sets, where n is any positive integer, we will derive a formula for the number of elements in the union of three sets A, B, and C. To construct this formula, we note that $|A| + |B| + |C|$ counts each element that is in exactly one of the three sets once, elements that are in exactly two of the sets twice, and elements in all three sets three times.

To remove the over-count of elements in more than one of the sets, we subtract the number of elements in the intersections of all pairs of the three sets. We obtain

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

This expression still counts elements that occur in exactly one of the sets once. An element that occurs in exactly two of the sets is also counted exactly once, because this element will occur in one of the three intersections of sets taken two at a time. However, those elements that occur in all three sets will be counted zero times by this expression, because they occur in all three intersections of sets taken two at a time.

To remedy this undercount, we add the number of elements in the intersection of all three sets. This final expression counts each element once, whether it is in one, two, or three of the sets. Thus,

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

Example 2. *A total of* 1232 *students have taken a course in Spanish,* 879 *have taken a course in French, and* 114 *have taken a course in Russian. Further,* 103 *have taken courses in both Spanish and French,* 23 *have taken courses in both Spanish and Russian, and* 14 *have taken courses in both French and Russian. If* 2092 *students have taken at least one of Spanish, French, and Russian, how many students have taken a course in all three languages?*

Solution. Let S be the set of students who have taken a course in Spanish, F the set of students who have taken a course in French, and R the set of students who have taken a course in Russian. Then

$$|S| = 1232, \quad |F| = 879, \quad |R| = 114, \quad |S \cap F| = 103, \quad |S \cap R| = 23, \quad |F \cap R| = 14,$$

and

$$|S \cup F \cup R| = 2092.$$

When we insert these quantities into the equation

$$|S \cup F \cup R| = |S| + |F| + |R| - |S \cap F| - |S \cap R| - |F \cap R| + |S \cap F \cap R|,$$

we obtain

$$2092 = 1232 + 879 + 114 - 103 - 23 - 14 + |S \cap F \cap R|.$$

We now solve for $|S \cap F \cap R|$. We find that $|S \cap F \cap R| = 7$. Therefore, there are seven students who have taken courses in Spanish, French, and Russian. □

Theorem 1 (THE PRINCIPLE OF INCLUSION-EXCLUSION)**.** *Let $A_1, A_2, \ldots, A_n$ be finite sets. Then*

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{1 \le i \le n} |A_i| - \sum_{1 \le i < j \le n} |A_i \cap A_n| + \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap \cdots \cap A_n|.$$

Proof. We will prove the formula by showing that an element in the union is counted exactly once by the right-hand side of the equation. Suppose that a is a member of exactly r of the sets $A_1, A_2, \ldots, A_n$, where $1 \le r \le n$. This element is counted $\binom{r}{1}$ times by $\sum |A_i|$. It is counted $\binom{r}{2}$ times by $\sum |A_i \cap A_j|$. In general, the element a is counted $\binom{r}{m}$ times by the summation involving m of the sets A_i. Thus, the element a is counted exactly

$$\binom{r}{1} - \binom{r}{2} + \binom{r}{3} - \cdots + (-1)^{r+1} \binom{r}{r}$$

times by the expression on the right-hand side of this equation. Our goal is to evaluate this quantity. By Corollary 2 of the binomial theorem in Section 6.4, we have

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

Hence,

$$1 = \binom{r}{0} = \binom{r}{1} - \binom{r}{2} + \binom{r}{3} - \cdots + (-1)^{r+1} \binom{r}{r}.$$

Therefore, each element in the union is counted exactly once by the expression on the right-hand side of the equation. This proves the principle of inclusion-exclusion. □