1.2.3 Cardinality: Countable and Uncountable Sets

Here we need to talk about **cardinality** of a set, which is basically the size of the set. The cardinality of a set is denoted by $|A|$. We first discuss cardinality for finite sets and then talk about infinite sets.

Finite Sets:
Consider a set A. If A has only a finite number of elements, its cardinality is simply the number of elements in A. For example, if $A = \{2, 4, 6, 8, 10\}$, then $|A| = 5$. Before discussing infinite sets, which is the main discussion of this section, we would like to talk about a very useful rule: the **inclusion-exclusion principle**. For two finite sets A and B, we have

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

To see this, note that when we add $|A|$ and $|B|$, we are counting the elements in $|A \cap B|$ twice, thus by subtracting it from $|A| + |B|$, we obtain the number of elements in $|A \cup B|$, (you can refer to Figure 1.16 in Problem 2 to see this pictorially). We can extend the same idea to three or more sets.

Inclusion-exclusion principle:

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

Generally, for n finite sets $A_1, A_2, A_3, \cdots, A_n$, we can write

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j|$$

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

Example 1.5
In a party,

- there are 10 people with white shirts and 8 people with red shirts;
- 4 people have black shoes and white shirts;
- 3 people have black shoes and red shirts;
- the total number of people with white or red shirts or black shoes is 21.

How many people have black shoes?

Solution

Let W, R, and B, be the number of people with white shirts, red shirts, and black shoes respectively. Then, here is the summary of the available information:

$$|W| = 10$$

$$|R| = 8$$

$$|W \cap B| = 4$$

$$|R \cap B| = 3$$

$$|W \cup B \cup R| = 21.$$

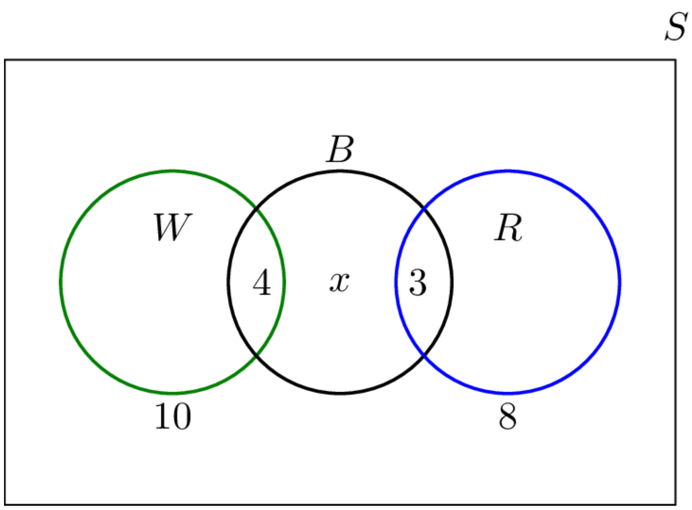
Also, it is reasonable to assume that W and R are disjoint, $|W \cap R| = 0$. Thus by applying the inclusion-exclusion principle we obtain

$$\begin{aligned} |W \cup R \cup B| &= 21 \\ &= |W| + |R| + |B| - |W \cap R| - |W \cap B| - |R \cap B| + |W \cap R \cap B| \\ &= 10 + 8 + |B| - 0 - 4 - 3 + 0. \end{aligned}$$

Thus

$$|B| = 10.$$

Note that another way to solve this problem is using a Venn diagram as shown in Figure 1.11.



$$21 = 10 + 8 + x \Rightarrow x = 3$$
$$\Rightarrow |B| = 4 + x + 3 = 10$$

Fig.1.11 - Inclusion-exclusion Venn diagram.

Infinite Sets:
What if A is an infinite set? It turns out we need to distinguish between two types of infinite sets, where one type is significantly "larger" than the other. In particular, one type is called **countable**, while the other is called **uncountable**. Sets such as $\mathbb{N}$ and $\mathbb{Z}$ are called countable, but "bigger" sets such as $\mathbb{R}$ are called uncountable. The difference between the two types is that you can list the elements of a countable set A, i.e., you can write $A = \{a_1, a_2, \cdots\}$, but you cannot list the elements in an uncountable set. For example, you can write

- $\mathbb{N} = \{1, 2, 3, \cdots\}$,
- $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \cdots\}$.

The fact that you can list the elements of a countably infinite set means that the set can be put in one-to-one correspondence with natural numbers $\mathbb{N}$. On the other hand, you cannot list the elements in $\mathbb{R}$, so it is an uncountable set. To be precise, here is the definition.

Definition 1.1
Set A is called countable if one of the following is true

a. if it is a finite set, $|A| < \infty$; or
b. it can be put in one-to-one correspondence with natural numbers $\mathbb{N}$, in which case the set is said to be countably infinite.
A set is called uncountable if it is not countable.

Here is a simple guideline for deciding whether a set is countable or not. As far as applied probability is concerned, this guideline should be sufficient for most cases.

- $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and any of their subsets are countable.
- Any set containing an interval on the real line such as $[a, b], (a, b], [a, b)$, or (a, b), where $a < b$ is uncountable.

The above rule is usually sufficient for the purpose of this book. However, to make the argument more concrete, here we provide some useful results that help us prove if a set is countable or not. If you are less interested in proofs, you may decide to skip them.

Theorem 1.3
Any subset of a countable set is countable.
Any superset of an uncountable set is uncountable.

Proof
The intuition behind this theorem is the following: If a set is countable, then any "smaller" set should also be countable, so a subset of a countable set should be countable as well. To provide a proof, we can argue in the following way.

Let A be a countable set and $B \subset A$. If A is a finite set, then $|B| \leq |A| < \infty$, thus B is countable. If A is countably infinite, then we can list the elements in A, then by removing the elements in the list that are not in B, we can obtain a list for B, thus B is countable.

The second part of the theorem can be proved using the first part. Assume B is uncountable. If $B \subset A$ and A is countable, by the first part of the theorem B is also a countable set which is a contradiction.

Theorem 1.4

If $A_1, A_2, \cdots$ is a list of countable sets, then the set $\bigcup_i A_i = A_1 \cup A_2 \cup A_3 \cdots$ is also countable.

Proof

It suffices to create a list of elements in $\bigcup_i A_i$. Since each A_i is countable we can list its elements: $A_i = \{a_{i1}, a_{i2}, \cdots\}$. Thus, we have

- $A_1 = \{a_{11}, a_{12}, \cdots\}$,
- $A_2 = \{a_{21}, a_{22}, \cdots\}$,
- $A_3 = \{a_{31}, a_{32}, \cdots\}$,
- ...

Now we need to make a list that contains all the above lists. This can be done in different ways. One way to do this is to use the ordering shown in Figure 1.12 to make a list. Here, we can write

$$\bigcup_i A_i = \{a_{11}, a_{12}, a_{21}, a_{31}, a_{22}, a_{13}, a_{14}, \cdots\} \tag{1.1}$$

$$
\begin{array}{ccccccccc}
a_{11} & \rightarrow & a_{12} & & a_{13} & \rightarrow & a_{14} & & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
a_{21} & & a_{22} & & a_{23} & & a_{24} & & \cdots \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & & \\
a_{31} & & a_{32} & & a_{33} & & a_{34} & & \cdots \\
 & \swarrow & & \nearrow & & & & & \\
a_{41} & & a_{42} & & a_{43} & & a_{44} & & \cdots \\
\downarrow & \nearrow & & & & & & & \\
a_{51} & & a_{52} & & a_{53} & & a_{54} & & \cdots \\
\vdots & & \vdots & & \vdots & & \vdots & &
\end{array}
$$

Fig.1.12 - Ordering to make a list.

We have been able to create a list that contains all the elements in $\bigcup_i A_i$, so this set is countable.

Theorem 1.5
If A and B are countable, then $A \times B$ is also countable.

Proof
The proof of this theorem is very similar to the previous theorem. Since A and B are countable, we can write

$$A = \{a_1, a_2, a_3, \cdots\},$$

$$B = \{b_1, b_2, b_3, \cdots\}.$$

Now, we create a list containing all elements in $A \times B = \{(a_i, b_j) | i, j = 1, 2, 3, \cdots\}$. The idea is exactly the same as before. Figure 1.13 shows one possible ordering.

$$
\begin{array}{ccccccccc}
(a_1,b_1) & \rightarrow & (a_1,b_2) & & (a_1,b_3) & \rightarrow & (a_1,b_4) & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \\
(a_2,b_1) & & (a_2,b_2) & & (a_2,b_3) & & (a_2,b_4) & \cdots \\
\downarrow & \nearrow & & \swarrow & & & & \\
(a_3,b_1) & & (a_3,b_2) & & (a_3,b_3) & & (a_3,b_4) & \cdots \\
 & \swarrow & & & & & & \\
(a_4,b_1) & & (a_4,b_2) & & (a_4,b_3) & & (a_4,b_4) & \cdots \\
\downarrow & & & & & & & \\
\vdots & & \vdots & & \vdots & & \vdots &
\end{array}
$$

Fig.1.13 - Ordering to make a list.

The above arguments can be repeated for any set C in the form of

$$C = \bigcup_{i} \bigcup_{j} \{a_{ij}\},$$

where indices i and j belong to some countable sets. Thus, any set in this form is countable. For example, a consequence of this is that the set of rational numbers $\mathbb{Q}$ is countable. This is because we can write

$$\mathbb{Q} = \bigcup_{i \in \mathbb{Z}} \bigcup_{j \in \mathbb{N}} \{\frac{i}{j}\}.$$

The above theorems confirm that sets such as $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and their subsets are countable. However, as we mentioned, intervals in $\mathbb{R}$ are uncountable. Thus, you can never provide a list in the form of $\{a_1, a_2, a_3, \cdots\}$ that contains all the elements in, say, $[0,1]$. This fact can be proved using a so-called diagonal argument, and we omit the proof here as it is not instrumental for the rest of the book.
