In general, the inclusion-exclusion formula can be written for any finite sequence:

\[ P \left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i) - \sum_{i_1 < i_2} P(A_{i_1} \cap A_{i_2}) + \sum_{i_1 < i_2 < i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n). \]

Of course, if the sets \(A_1, A_2, \ldots, A_n\) are pairwise disjoint, then we arrive back at

\[ P \left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i), \]

a result implied by Axiom 3 by taking \(A_{n+1} = A_{n+2} = \cdots = \emptyset\).

### 2.5 Discrete probability models and events

**TERMINOLOGY:** If a sample space for an experiment contains a finite or countable number of sample points, we call it a **discrete sample space**.

- **Finite:** “number of sample points < \(\infty\).”
- **Countable:** “number of sample points may equal \(\infty\), but can be counted; i.e., sample points may be put into a 1:1 correspondence with \(\mathcal{N} = \{1, 2, \ldots\}\).”

**Example 2.7.** A standard roulette wheel contains an array of numbered compartments referred to as “pockets.” The pockets are either red, black, or green. The numbers 1 through 36 are evenly split between red and black, while 0 and 00 are green pockets. On the next play, we are interested in the following events:

\[ A_1 = \{13\} \]
\[ A_2 = \{\text{red}\} \]
\[ A_3 = \{0, 00\}. \]

**TERMINOLOGY:** A **simple event** is an event that can not be decomposed. That is, a simple event corresponds to exactly one sample point. **Compound events** are those events that contain more than one sample point. In Example 2.7, because \(A_1\) contains...
only one sample point, it is a simple event. The events \( A_2 \) and \( A_3 \) contain more than one sample point; thus, they are compound events.

**STRATEGY:** Computing the probability of a compound event can be done by

1. counting up all sample points associated with the event (this can be very easy or very difficult)
   
   \[ P(A_3) = P(\{0.00\}) = P(\{0.04\}) + P(\{0.03\}) \]

2. adding up the probabilities associated with each sample point.

**NOTATION:** Your authors use the symbol \( E_i \) to denote the \( i \)th sample point (i.e., \( i \)th simple event). Thus, adopting the aforementioned strategy, if \( A \) denotes any compound event,

\[
P(A) = \sum_{E_i \in A} P(E_i).
\]

We simply sum up the simple event probabilities \( P(E_i) \) for all \( i \) such that \( E_i \in A \).

**Example 2.8.** *An equiprobability model.* Suppose that a discrete sample space \( S \) contains \( N < \infty \) sample points, each of which are equally likely. If the event \( A \) consists of \( n_a \) sample points, then \( P(A) = n_a/N \).

**Proof.** Write \( S = E_1 \cup E_2 \cup \cdots \cup E_N \), where \( E_i \) corresponds to the \( i \)th sample point; \( i = 1, 2, \ldots, N \). Then,

\[
1 = P(S) = P(E_1 \cup E_2 \cup \cdots \cup E_N) = \sum_{i=1}^{N} P(E_i).
\]

Now, as \( P(E_1) = P(E_2) = \cdots = P(E_N) \), we have that

\[
1 = \sum_{i=1}^{N} P(E_i) = NP(E_1),
\]

and, thus, \( P(E_1) = \frac{1}{N} = P(E_2) = \cdots = P(E_N) \). Without loss of generality, take \( A = E_1 \cup E_2 \cup \cdots \cup E_{n_a} \). Then,

\[
P(A) = P(E_1 \cup E_2 \cup \cdots \cup E_{n_a}) = \sum_{i=1}^{n_a} P(E_i) = \sum_{i=1}^{n_a} \frac{1}{N} = n_a/N. \quad \Box
\]
Example 2.9. Two jurors are needed from a pool of 2 men and 2 women. The jurors are randomly selected from the 4 individuals. A sample space for this experiment is

\[ S = \{(M1, M2), (M1, W1), (M1, W2), (M2, W1), (M2, W2), (W1, W2)\}. \]

What is the probability that the two jurors chosen consist of 1 male and 1 female?

**Solution.** There are \( N = 6 \) sample points, denoted in order by \( E_1, E_2, ..., E_6 \). Let the event

\[ A = \{ \text{one male, one female} \} = \{(M1, W1), (M1, W2), (M2, W1), (M2, W2)\}, \]

so that \( n_A = 4 \). If the sample points are equally likely (probably true if the jurors are randomly selected), then \( P(A) = \frac{4}{6} \). \( \square \)

2.6 Tools for counting sample points

2.6.1 The multiplication rule

**MULTIPLICATION RULE:** Consider an experiment consisting of \( k \geq 2 \) “stages,” where

\[ n_1 = \text{number of ways stage 1 can occur} \]
\[ n_2 = \text{number of ways stage 2 can occur} \]
\[ \vdots \]
\[ n_k = \text{number of ways stage } k \text{ can occur}. \]

Then, there are

\[ \prod_{i=1}^{k} n_i = n_1 \times n_2 \times \cdots \times n_k \]

different outcomes in the experiment.

Example 2.10. An experiment consists of rolling two dice. Envision stage 1 as rolling the first and stage 2 as rolling the second. Here, \( n_1 = 6 \) and \( n_2 = 6 \). By the multiplication rule, there are \( n_1 \times n_2 = 6 \times 6 = 36 \) different outcomes. \( \square \)