1.3.2 Probability

We assign a **probability** measure $P(A)$ to an event A. This is a value between 0 and 1 that shows how likely the event is. If $P(A)$ is close to 0, it is very unlikely that the event A occurs. On the other hand, if $P(A)$ is close to 1, A is very likely to occur. The main subject of probability theory is to develop tools and techniques to calculate probabilities of different events. Probability theory is based on some axioms that act as the foundation for the theory, so let us state and explain these axioms.

Axioms of Probability:

- Axiom 1: For any event A, $P(A) \geq 0$.

- Axiom 2: Probability of the sample space S is $P(S) = 1$.

- Axiom 3: If $A_1, A_2, A_3, \cdots$ are disjoint events, then
  $P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$

Let us take a few moments and make sure we understand each axiom thoroughly. The first axiom states that probability cannot be negative. The smallest value for $P(A)$ is zero and if $P(A) = 0$, then the event A will never happen. The second axiom states that the probability of the whole sample space is equal to one, i.e., 100 percent. The reason for this is that the sample space S contains all possible outcomes of our random experiment. Thus, the outcome of each trial always belongs to S, i.e., the event S always occurs and $P(S) = 1$. In the example of rolling a die, $S = \{1, 2, 3, 4, 5, 6\}$, and since the outcome is always among the numbers 1 through 6, $P(S) = 1$.

The third axiom is probably the most interesting one. The basic idea is that if some events are disjoint (i.e., there is no overlap between them), then the probability of their union must be the summations of their probabilities. Another way to think about this is to imagine the probability of a set as the area of that set in the Venn diagram. If

several sets are disjoint such as the ones shown [Figure 1.9](), then the total area of their union is the sum of individual areas. The following example illustrates the idea behind the third axiom.

Example 1.8

In a presidential election, there are four candidates. Call them A, B, C, and D. Based on our polling analysis, we estimate that A has a 20 percent chance of winning the election, while B has a 40 percent chance of winning. What is the probability that A **or** B win the election?

Solution

Notice that the events that $\{A \text{ wins}\}$, $\{B \text{ wins}\}$, $\{C \text{ wins}\}$, and $\{D \text{ wins}\}$ are disjoint since more than one of them cannot occur at the same time. For example, if A wins, then B cannot win. From the third axiom of probability, the probability of the union of two disjoint events is the summation of individual probabilities. Therefore,

$$\begin{aligned} P(\text{A wins or B wins}) &= P\big(\{A \text{ wins}\} \cup \{B \text{ wins}\}\big) \\ &= P(\{A \text{ wins}\}) + P(\{B \text{ wins}\}) \\ &= 0.2 + 0.4 \\ &= 0.6 \end{aligned}$$

In summary, if A_1 and A_2 are disjoint events, then $P(A_1 \cup A_2) = P(A_1) + P(A_2)$. The same argument is true when you have n disjoint events $A_1, A_2, \cdots, A_n$:

$$P(A_1 \cup A_2 \cup A_3 \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n), \text{ if } A_1, A_2, \cdots, A_n \text{ are disjoint.}$$

In fact, the third axiom goes beyond that and states that the same is true even for a countably infinite number of disjoint events. We will see more examples of how we use the third axiom shortly.

As we have seen, when working with events, *intersection* means "*and*", and *union* means "*or*". The probability of intersection of A and B, $P(A \cap B)$, is sometimes shown by $P(A, B)$ or $P(AB)$.

Notation:

- $P(A \cap B) = P(A \text{ and } B) = P(A, B)$,

- $P(A \cup B) = P(A \text{ or } B)$.