Sample space and events A sample space Ω is set of all possible outcomes (elementary events). An element of Ω is an elementary event and is denoted ω. An event E is a subset of possible outcomes. Example: rolling a die, $\Omega = \{1, 2, 3, 4, 5, 6\}$ and an example of E is rolling an even number.

Probability A probability function P is a function mapping events to real numbers that satisfies these conditions:

1. for any event E, $0 \leq P(E) \leq 1$
2. $P(\Omega) = 1$ (one elementary event must happen)
3. for any set of disjoint events $E_1, E_2, \ldots, E_n$, $P(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{i=1}^{n} P(E_i)$

For any two events, $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

Union bound $P(E_1 \cup E_2 \cup \cdots \cup E_n) \leq \sum_{i=1}^{n} P(E_i)$

Conditional probability The conditional probability that event A occurs given that B occurs is

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Of course, for this to make sense, it must be that $P(A) > 0$.

Chain rule For any two events A and B, $P(A \cap B) = P(B|A)P(A)$. It's also true that $P(A \cap B) = P(A|B)P(B)$. (Follows from definition of conditional probability and rearranging terms.)

Independent events Events A and B are independent if $P(A \cap B) = P(A)P(B)$. Intuitively events are independent if the occurrence of one event does not affect the probability that the other occurs. Independence implies $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

Random variables

An event is simply a description of a possible outcome (e.g., "I roll a 4", "I am struck by a meteor"). Descriptions are hard to work with. By mapping events to numbers, we can apply mathematical operations to them.

Random variable Formally, a random variable X is a function that maps events to numbers. We can use the notation $X(\omega) = x$ to mean that event ω is mapped to x. The probability that X takes on the value x is simply: $P(X = x) = \sum_{\omega \in \Omega : X(\omega) = x} P(\omega)$.

Expected value Suppose X takes on n possible values $x_1, \ldots, x_n$. The expected value of a random variable is $E[X] = \sum_{i=1}^{n} x_i P(X = x_i)$.

Marginalization / Total law of probability Let X and Y be two random variables that take on the values $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ respectively.

$$P(X = x_i) = \sum_{j=1}^{m} P(X = x_i \cap Y = y_j)$$

Linearity of Expectations Given random variables X and Y, $E[X + Y] = E[X] + E[Y]$.

$$\begin{aligned}
E[X+Y] &= \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i + y_j)P(X = x_i \cap Y = y_j) \\
&= \sum_{i=1}^{n} x_i \sum_{j=1}^{m} P(X = x_i \cap Y = y_j) + \sum_{j=1}^{m} y_j \sum_{i=1}^{n} P(X = x_i \cap Y = y_j) \\
&= \sum_{i=1}^{n} x_i P(X = x_i) + \sum_{j=1}^{m} y_j P(Y = y_j) = E[X] + E[Y]
\end{aligned}$$

Bayes' rule This simple but powerful rule follows from the definition of conditional probability and an application of the chain rule.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Let $\neg A$ denote the event "A did not occur," we can also write it as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)}$$

A particular useful application is when event A is some hidden cause and event B is some observable symptom. Given the test result is positive (B), what is the probability the patient has the disease (A)?