# A Primer on Probability Theory for Algorithms

## 1 Probability Spaces and Events

A *discrete probability space* is a pair $(\Omega, \mathbf{Pr})$, where

- $\Omega$: the sample space, which is the set of all possible outcomes.
- $\mathbf{Pr}$: *probability distribution*, which is a function $\mathbf{Pr} : \Omega \to \mathbb{R}_{\geq 0}$ satisfying

$$\sum_{\omega \in \Omega} \mathbf{Pr}(\omega) = 1$$

Every subset $\mathcal{E}$ of $\Omega$ is called an *event*.

$$\mathbf{Pr}(\mathcal{E}) := \sum_{\omega \in \mathcal{E}} \mathbf{Pr}(\omega)$$

For example, in a coin flip experiment, $\Omega = \{H, T\}$, where $H$ denotes heads and $T$ denotes tails. An event might be the set $\mathcal{E} = \{H\}$, meaning heads occur, with probability $\mathbf{Pr}(\{H\}) = 0.5$.

## 2 Random Variables

A *random variable* is a function $X : \Omega \to \mathbb{R}$, which assigns a real number to each outcome in $\Omega$. For example, in rolling a die, the random variable $X$ might represent the number that comes up. We often denote the probability that $X$ takes on a specific value $x$ as $\mathbf{Pr}(X = x)$.

## 3 Expectation and Linearity of Expectation

The *expected value* or *expectation* of a random variable $X$, denoted $\mathbb{E}[X]$, is the average value $X$ would take if the experiment were repeated infinitely many times. For a discrete random variable:

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} x \cdot \mathbf{Pr}(X = x)$$

An important property of expectation is its linearity. For any two random variables $X$ and $Y$ and any constants $a$ and $b$, we have:

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

This property holds regardless of whether $X$ and $Y$ are independent, making it a powerful tool in analyzing algorithms.

# 4 Distributions

## 4.1 Uniform Distribution

The *uniform distribution* is a simple probability distribution where all outcomes are equally likely. If $X$ is a random variable representing the roll of a fair die, then $\mathbf{Pr}(X = k) = \frac{1}{6}$ for $k \in \{1, 2, 3, 4, 5, 6\}$.

## 4.2 Geometric Distribution

The *geometric distribution* models the number of trials until the first success in a sequence of Bernoulli trials. If the probability of success on each trial is $p$, then the probability that the first success occurs on the $k$-th trial is given by:

$$\mathbf{Pr}(X = k) = (1 - p)^{k-1} p$$

The expected value of a geometrically distributed random variable is

$$\mathbb{E}[X] = \frac{1}{p}$$