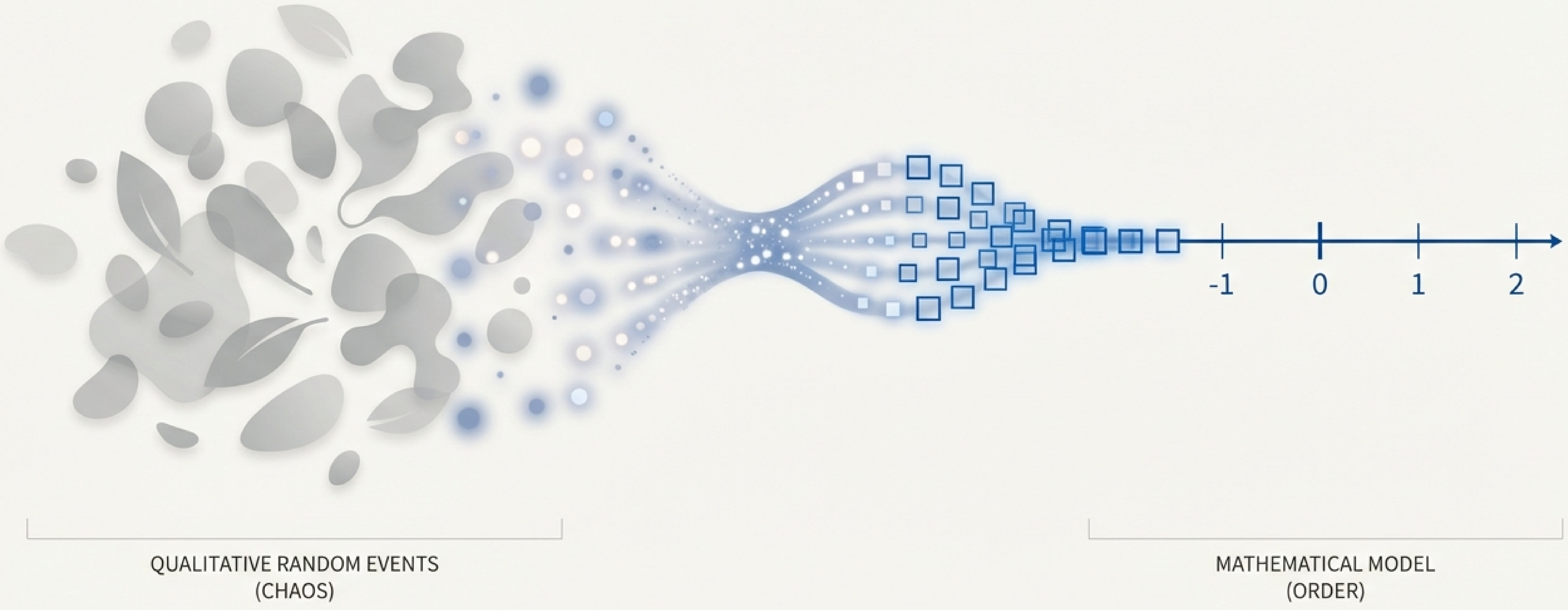
RANDOM VARIABLES: THE LANGUAGE OF UNCERTAINTY

Translating Random Events into Mathematical Models



NotebookLM

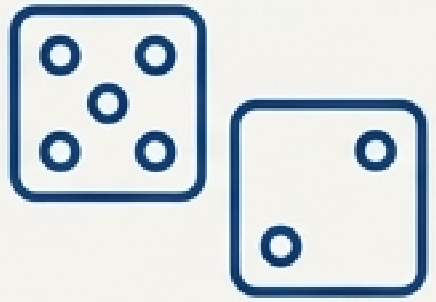
From Chance to Computation

How do we mathematically analyze outcomes that aren’t numbers?

Tossing a coin

Outcome: “Heads” or “Tails”

Rolling two dice



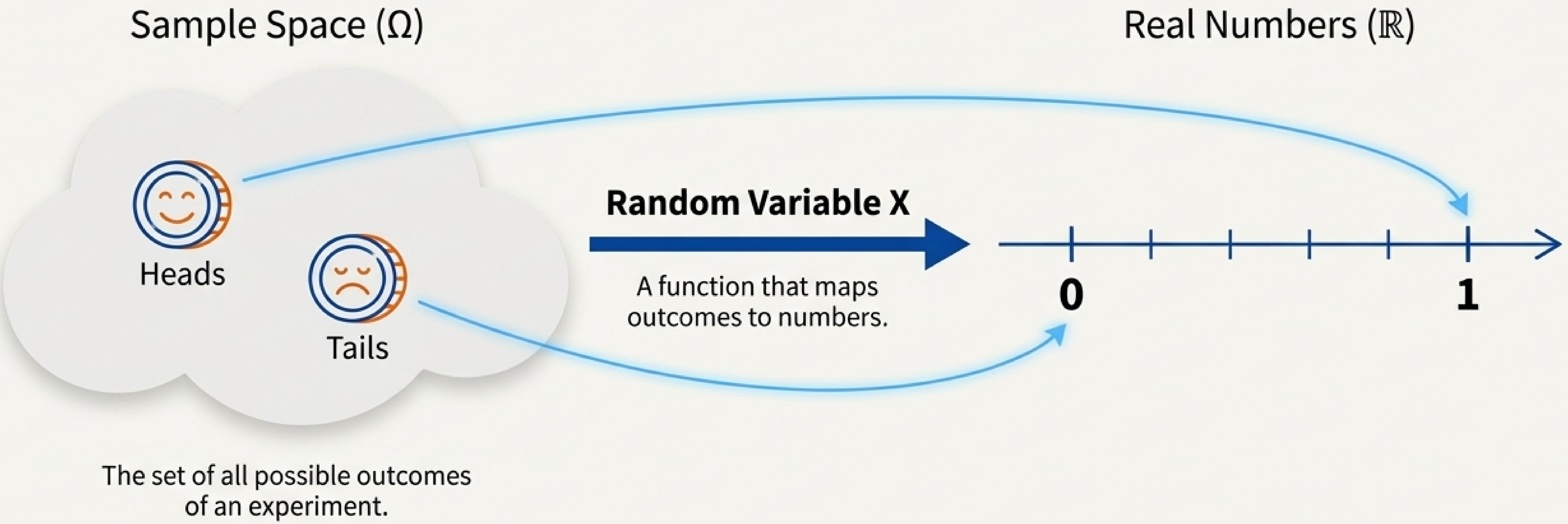
Outcome: A pair of faces

Measuring wait time

Outcome: “Some duration of time”

We often describe events with words. To analyze them mathematically, we need to think of them as numbers. This requires a bridge from the qualitative to the quantitative.

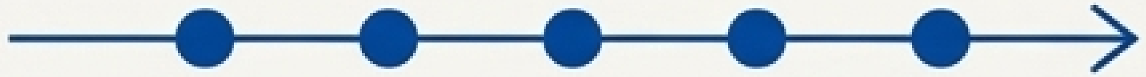
The Random Variable: A Bridge from Events to Numbers



A random variable is a function that assigns a numerical value to the outcome of a random event. Given a sample space Ω, a random variable X is a function $X : \Omega \to \mathbb{R}$ that maps every element ω of the sample space to a value $X(\omega) \in \mathbb{R}$.

Discrete vs. Continuous: Counting vs. Measuring

Discrete Random Variables

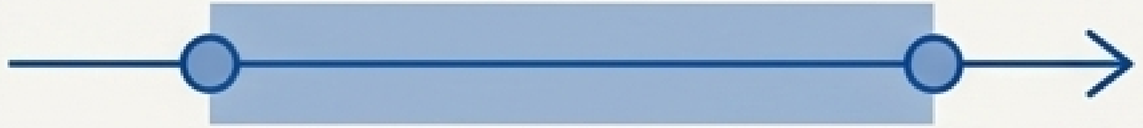


A discrete random variable takes on a *countable* number of possible values. A countable set is either finite or can be put in one-to-one correspondence with the set of natural numbers.

Examples:

- **Number of heads in 3 coin flips** (possible values: 0, 1, 2, 3).
- **Sum of two dice rolls** (possible values: 2, 3, ..., 12).

Continuous Random Variables



A continuous random variable takes on an infinite, *uncountable* number of possible values within a given range.

Examples:

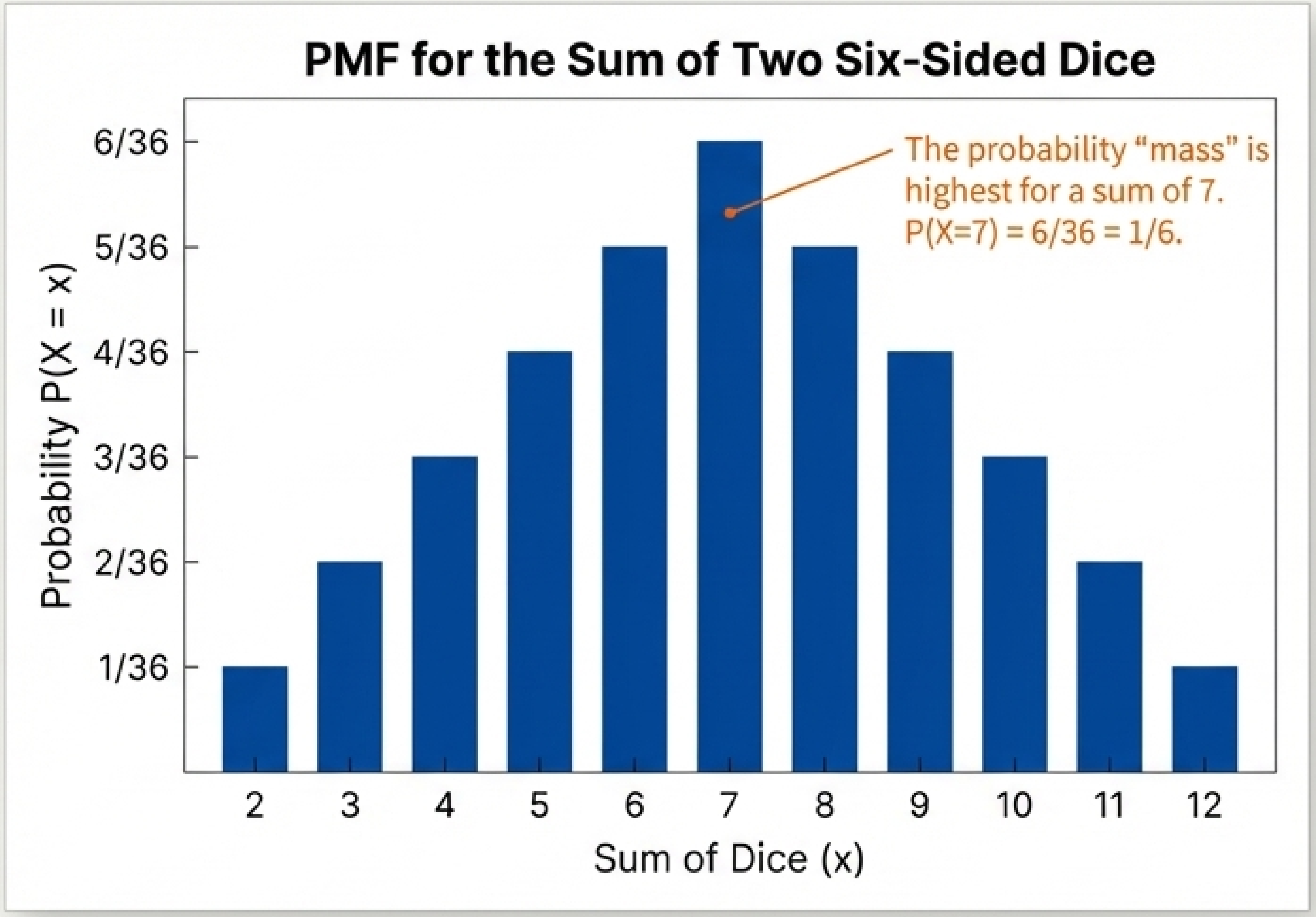
- The **exact waiting time for a train** (e.g., any value in $[0, \infty)$ minutes).
- The **precise height of a randomly selected student** (e.g., any value in $[150, 200]$ cm).

Describing the Discrete: The Probability Mass Function (PMF)

Pinpointing Probability for Discrete Variables

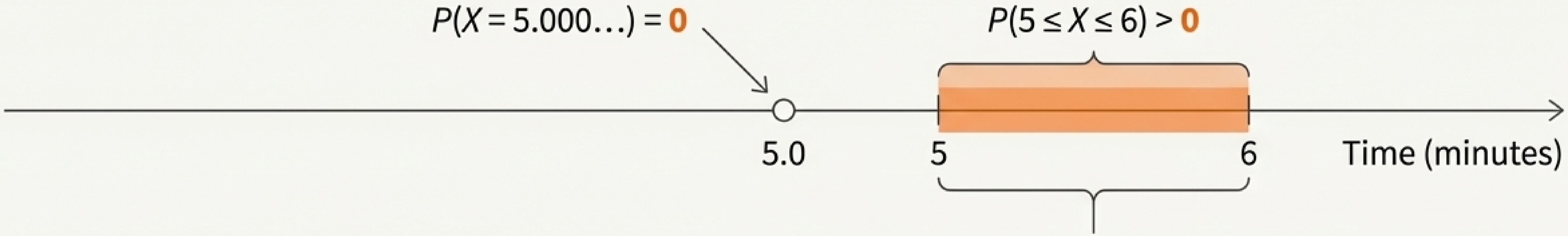
The probability mass function (PMF) of a discrete random variable X is a function that gives the probability that X will take on a specific value. The PMF is sometimes called the discrete probability distribution.

$$p_X(x) = P(X = x)$$



The Paradox of Infinite Precision

For a continuous random variable X, what is the probability of an exact value, e.g., $P(X = 27.14159\ldots)$?



A continuous variable can take on an infinite, uncountable number of values within any range.

Therefore, the probability of hitting any one *exact* value is effectively zero.

Insight: With a continuous random variable X, we usually look at the probability that X falls in a certain range of values, rather than being equal to a single value.

Example: It makes no sense to ask for the probability of a train arriving in *exactly* 5.000… minutes. It makes sense to examine the probability over an interval, for example $P(5 \le X \le 6)$.

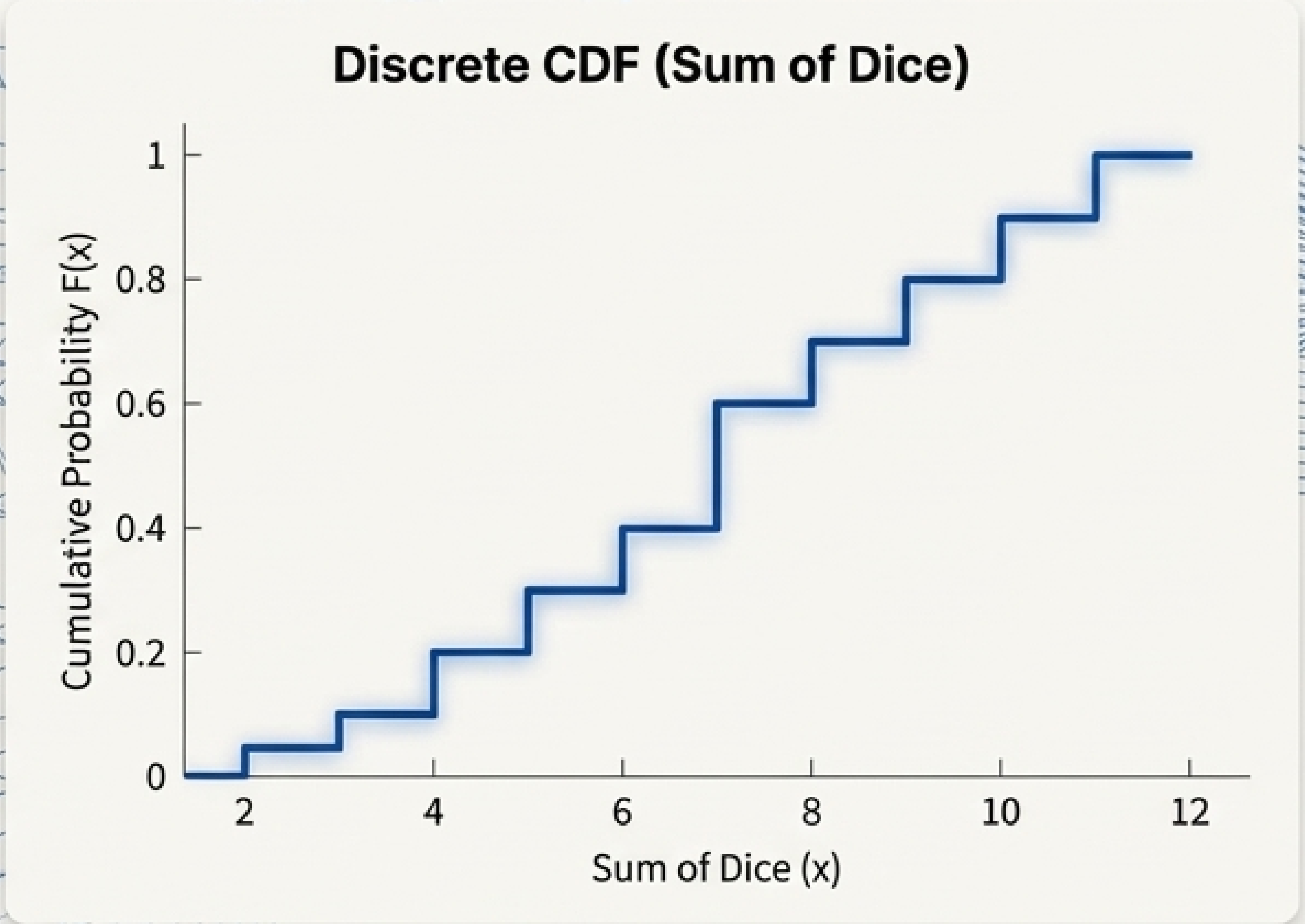
NotebookLM

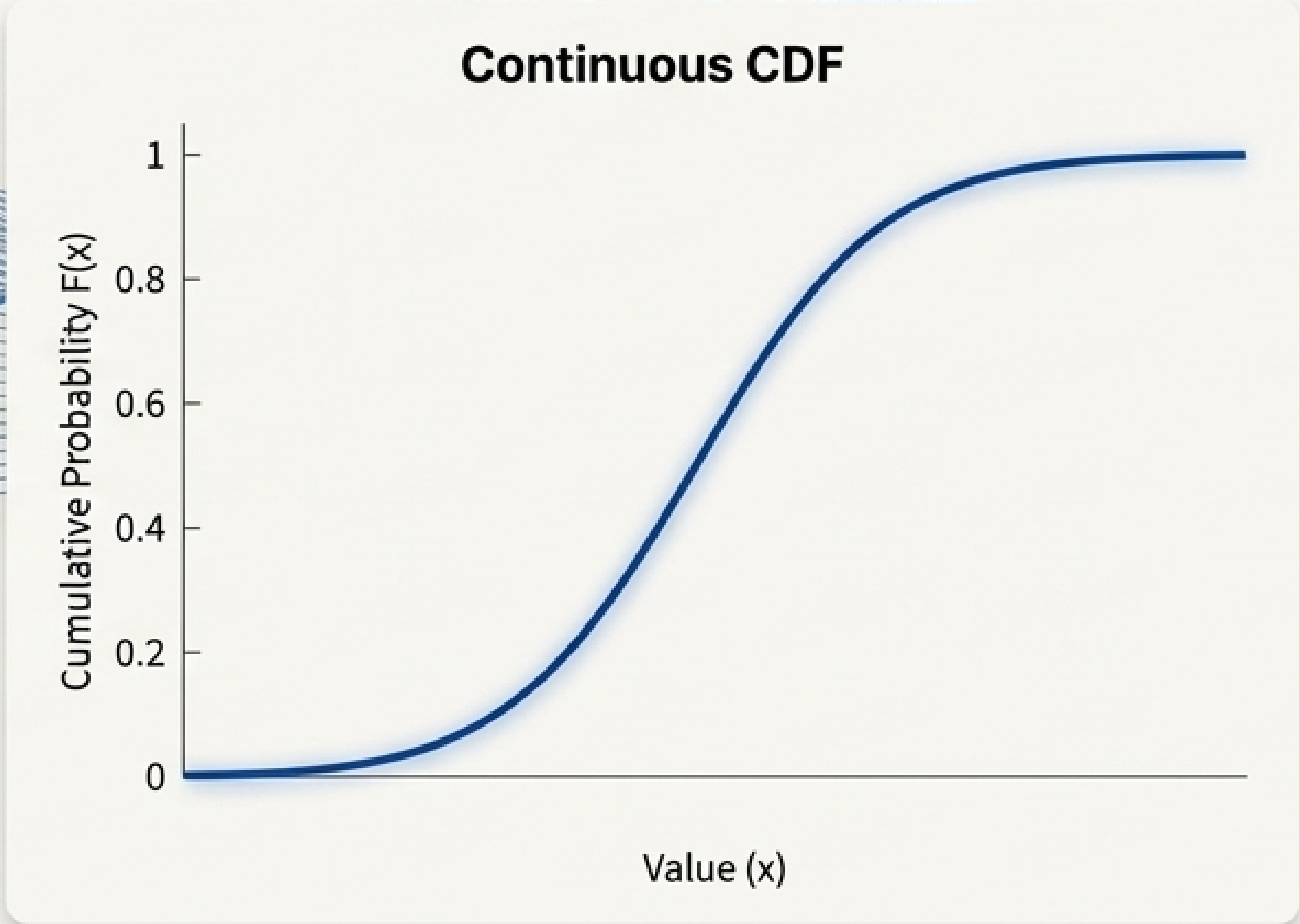
A Universal Language: The Cumulative Distribution Function (CDF)

The One Function That Describes All Random Variables

The cumulative distribution function (CDF) of a random variable X is a function that gives the probability that X will take on a value less than or equal to some number $x \in \mathbb{R}$. The CDF is defined for both discrete and continuous random variables.

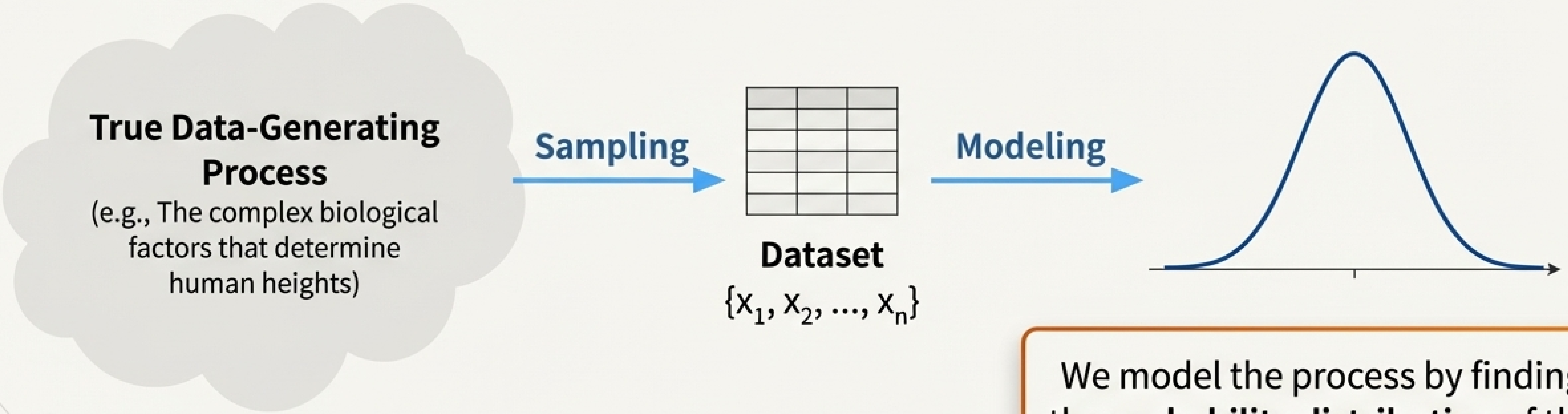
$$F_X(x) = P(X \leq x)$$





Why This is the Bedrock of Machine Learning

In machine learning, we treat our dataset as a collection of **realizations** (observed values) of one or more random variables. Our goal is to model the underlying data-generating process.



We model the process by finding the **probability distribution** of the underlying Random Variable.

Summary: The Core Ideas

The Bridge

A Random Variable is a **function** ($X: \Omega \to \mathbb{R}$) that maps random outcomes to real numbers, enabling mathematical analysis.

The Two Forms

Discrete variables involve a *countable* set of values. **Continuous** variables can take any value in an *uncountable* range.

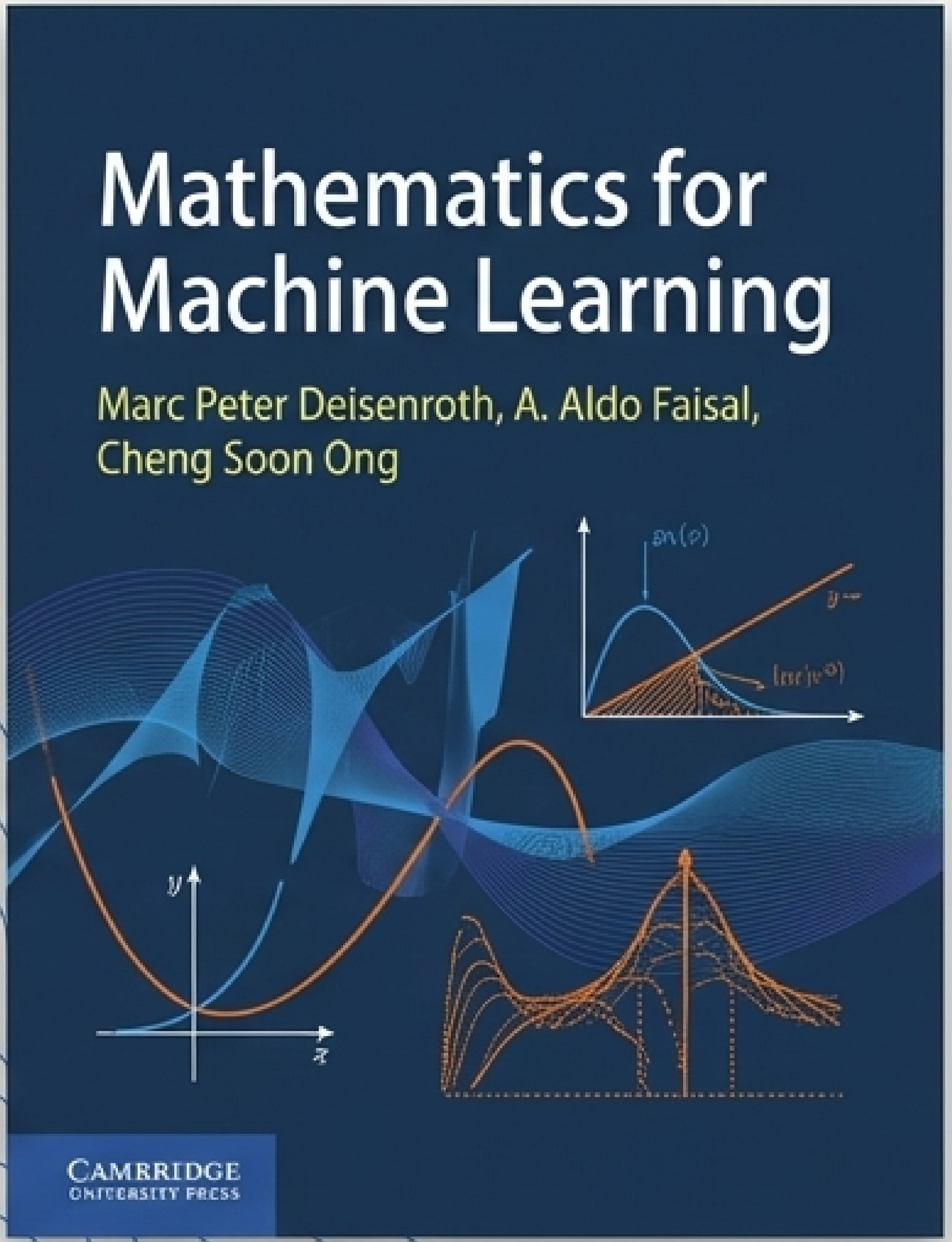
The Right Tools

For discrete variables, the **PMF** gives $P(X = x)$.
For all random variables, the **CDF** gives the cumulative probability $P(X \leq x)$.

The ML Context

Data points are samples (realizations) of random variables. Machine learning models are often attempts to learn their underlying distributions.

Continue Your Journey



This presentation is based on concepts from **Chapter 6**: **Probability and Distributions**.

This self-contained textbook bridges the gap between mathematical and machine learning texts, introducing the mathematical concepts with a minimum of prerequisites.

https://mml-book.com

The Language of Uncertainty