



UNIVERSITY OF
TORONTO

Statistical Sciences

DoSS Summer Bootcamp Probability Module 1

Ichiro Hashimoto

University of Toronto

July 7, 2026

Roadmap

A bridge connecting undergraduate probability and graduate probability

Undergraduate-level probability

- Concrete;
- Examples and scenarios;
- Rely on computation...

Roadmap

A bridge connecting undergraduate probability and graduate probability

Undergraduate-level probability

- Concrete;
- Examples and scenarios;
- Rely on computation...

Graduate-level probability

- Abstract (measure theory);
- Laws and properties;
- Rely on construction and inference...

Roadmap

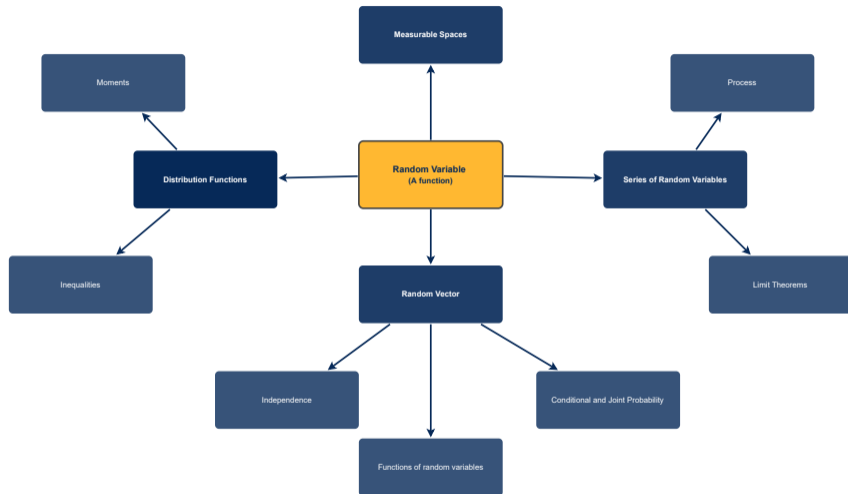


Figure: Roadmap

Outline

- Measurable spaces
 - ▷ Sample Space
 - ▷ σ -algebra
- Probability measures
 - ▷ Measures on σ -field
 - ▷ Basic results
- Conditional probability
 - ▷ Bayes' rule
 - ▷ Law of total probability

Measurable spaces

Sample Space

The sample space Ω is the set of all possible outcomes of an experiment.

Examples:

- Toss a coin: $\{H, T\}$
- Roll a die: $\{1, 2, 3, 4, 5, 6\}$

Measurable spaces

Sample Space

The sample space Ω is the set of all possible outcomes of an experiment.

Examples:

- Toss a coin: $\{H, T\}$
- Roll a die: $\{1, 2, 3, 4, 5, 6\}$

Event

An event is a collection of possible outcomes (subset of the sample space).

Examples:

- Get head when tossing a coin: $\{H\}$
- Get an even number when rolling a die: $\{2, 4, 6\}$

Discrete and Continuous Probability

Discrete random variable

For a discrete random variable X ,

$$\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x).$$

The probability of an event is obtained by adding point probabilities.

Continuous random variable

For a continuous random variable X with density f ,

$$\mathbb{P}(X \in A) = \int_A f(x) dx.$$

The probability of an event is obtained by integrating the density.

Why do we have two different definitions?

Discrete and Continuous Expectations

Discrete random variable

For a discrete random variable X ,

$$\mathbb{E}[X] = \sum_{x \in \mathbb{R}} x \mathbb{P}(X = x).$$

Continuous random variable

For a continuous random variable X with density f ,

$$\mathbb{E}[X] = \int_{\mathbb{R}} xf(x) dx.$$

Again, a **sum** is used under discrete probability, while an **integral** is used under continuous probability.

Why Can't We Always Use Sums?

Suppose X is a continuous random variable. Then

$$\mathbb{P}(X = x) = 0 \quad \text{for every } x \in \mathbb{R}.$$

If we imitate discrete probability, then

$$1 = \mathbb{P}(X \in \mathbb{R}) = \sum_{x \in \mathbb{R}} \mathbb{P}(X = x) = \sum_{x \in \mathbb{R}} 0.$$

This would suggest

$$1 = 0,$$

which is an obvious contradiction.

Key point

The problem is not that each point has probability zero.

The problem is that we are trying to add **uncountably many** numbers.



Countability Is the Key

In ordinary calculus, infinite sums are defined for **countable** values:

$$\sum_{i=1}^{\infty} a_i.$$

But an expression such as

$$\sum_{x \in [0,1]} a_x$$

is not a countable sum but an **uncountable** sum.

Motivation

Measure-theoretic probability is built around **countable** operations.
This motivates the definition of a σ -**algebra**.

A Countable Approximation of the Real Line

For each positive integer n , divide the real line into countably many intervals:

$$\left[\frac{i}{n}, \frac{i+1}{n} \right), \quad i \in \mathbb{Z}.$$

Then

$$\mathbb{R} = \bigcup_{i=-\infty}^{\infty} \left[\frac{i}{n}, \frac{i+1}{n} \right).$$

These intervals are disjoint, so countable additivity gives

$$\mathbb{P}(X \in \mathbb{R}) = \sum_{i=-\infty}^{\infty} \mathbb{P} \left(X \in \left[\frac{i}{n}, \frac{i+1}{n} \right) \right).$$

Key point

Even though \mathbb{R} is uncountable, it can be decomposed into **countably many** intervals.



Approximating Expectation by Countable Sums

Using the same intervals,

$$\left[\frac{i}{n}, \frac{i+1}{n} \right), \quad i \in \mathbb{Z},$$

we approximate X by the left endpoint of the interval containing it.

That is,

$$X \approx \frac{i}{n} \quad \text{when } X \in \left[\frac{i}{n}, \frac{i+1}{n} \right).$$

Therefore,

$$\mathbb{E}[X] \approx \sum_{i=-\infty}^{\infty} \frac{i}{n} \mathbb{P} \left(X \in \left[\frac{i}{n}, \frac{i+1}{n} \right) \right).$$

Important observation

This formula uses only a **countable sum**, not an uncountable sum over individual points.

From Approximation to Measure-Theoretic Probability

As n becomes larger, the intervals

$$\left[\frac{i}{n}, \frac{i+1}{n} \right)$$

become smaller.

So the approximation should become more accurate

$$\mathbb{E}[X] \stackrel{?}{=} \lim_{n \rightarrow \infty} \sum_{i=-\infty}^{\infty} \frac{i}{n} \mathbb{P} \left(X \in \left[\frac{i}{n}, \frac{i+1}{n} \right) \right).$$

Key insights

These observations suggest that probabilities and expectations can be built from:

- probabilities of sets,
- countable unions of sets,
- countable sums.

Main idea

Measure-theoretic probability gives one unified language for both discrete and continuous probability by focusing on **sets** and **countable operations**.

Introduction to Measure Theory

σ -algebra

A σ -algebra (σ -field) \mathcal{F} on Ω is a non-empty collection of subsets of Ω such that

- If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$,
- If $A_1, A_2, \dots \in \mathcal{F}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Remark: $\emptyset, \Omega \in \mathcal{F}$

Probability measures

Measures on σ -field

A function $\mu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ is called a measure if

- $\mu(\emptyset) = 0$,
- If $A_1, A_2, \dots \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$, then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

If $\mu(\Omega) = 1$, then μ is called a probability measure.

Probability measures

Measures on σ -field

A function $\mu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ is called a measure if

- $\mu(\emptyset) = 0$,
- If $A_1, A_2, \dots \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$, then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

If $\mu(\Omega) = 1$, then μ is called a probability measure.

Properties:

- Monotonicity: $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$
- Subadditivity: $A \subseteq \cup_{i=1}^{\infty} A_i \Rightarrow \mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$
- Continuity from below: $A_i \nearrow A \Rightarrow \mu(A_i) \nearrow \mu(A)$
- Continuity from above: $A_i \searrow A$ and $\mu(A_i) < \infty \Rightarrow \mu(A_i) \searrow \mu(A)$

Probability measures

Proof of continuity from below:

Probability measures

Proof of continuity from above:

Remark: $\mu(A_i) < \infty$ is vital.

Example: Tossing a Coin Twice

Consider the sample space $\Omega = \{HH, HT, TH, TT\}$.

One possible σ -algebra is the power set

$$2^\Omega,$$

which contains all $2^4 = 16$ subsets of Ω .

Another example is

$$\mathcal{F} = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}.$$

This collection satisfies

- $\emptyset, \Omega \in \mathcal{F}$,
- complements remain in \mathcal{F} ,
- countable unions remain in \mathcal{F} .

The Smallest and Largest σ -Algebras

For any sample space Ω , there are always two extreme examples.

The trivial σ -algebra

$$\{\emptyset, \Omega\}.$$

Only two events are measurable.

The power set

$$2^\Omega = \{A : A \subseteq \Omega\}.$$

Every subset of Ω is measurable.

Observation

Every σ -algebra lies between these two extremes:

$$\{\emptyset, \Omega\} \subseteq \mathcal{F} \subseteq 2^\Omega.$$

Generating a σ -Algebra

Instead of specifying every measurable set, we often begin with a collection of subsets.
Suppose

$$\mathcal{C} = \{A_1, A_2, \dots\}.$$

We then define $\sigma(\mathcal{C})$ to be the **smallest σ -algebra containing every set in \mathcal{C}** .

Intuitively, $\sigma(\mathcal{C})$

- include every set in \mathcal{C} ,
- include all complements,
- include all countable unions,
- repeat until no new sets are created.

This construction allows us to build useful σ -algebras from simple building blocks.

Example: The Borel σ -Algebra

On the real line, the most important σ -algebra is the **Borel σ -algebra**. It is generated by all open intervals:

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(a, b) : a < b\}).$$

Therefore,

$$(a, b) \in \mathcal{B}(\mathbb{R})$$

for every open interval, and so are all sets obtained by repeatedly taking

- complements,
- countable unions,
- countable intersections.

Examples of Borel sets

Examples of Borel sets include

$[a, b]$, (a, b) , $(-\infty, a]$, \mathbb{Z} , \mathbb{Q} .

Conditional probability

Original problem:

- What is the probability of some event A ?
- $P(A)$ is determined by our probability measure.

New problem:

- Given that B happens, what is the probability of some event A ?
- $P(A | B)$ is the conditional probability of the event A given B .

Conditional probability

Original problem:

- What is the probability of some event A ?
- $P(A)$ is determined by our probability measure.

New problem:

- Given that B happens, what is the probability of some event A ?
- $P(A | B)$ is the conditional probability of the event A given B .

Example:

- Roll a die: $P(\{2\} | \text{even number})$

Conditional probability

Definition

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

Remark: Does conditional probability $P(\cdot | B)$ satisfy the axioms of a probability measure?

Conditional probability

Multiplication rule

$$P(A \cap B) = P(A | B)P(B) = P(B | A)P(A)$$

Generalization:

Law of total probability

Let A_1, A_2, \dots, A_n be a partition of ω , such that $P(A_i) > 0$, then

$$P(B) = \sum_{i=1}^n P(A_i)P(B | A_i)$$

Problem Set

Problem 1: Prove that for a σ -algebra \mathcal{F} , if $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Problem 2: Prove monotonicity and subadditivity of measure μ on σ -algebra.

Problem 3: (Monty Hall problem) Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

(Assumptions: the host will not open the door we picked and the host will only open the door which has a goat.)