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

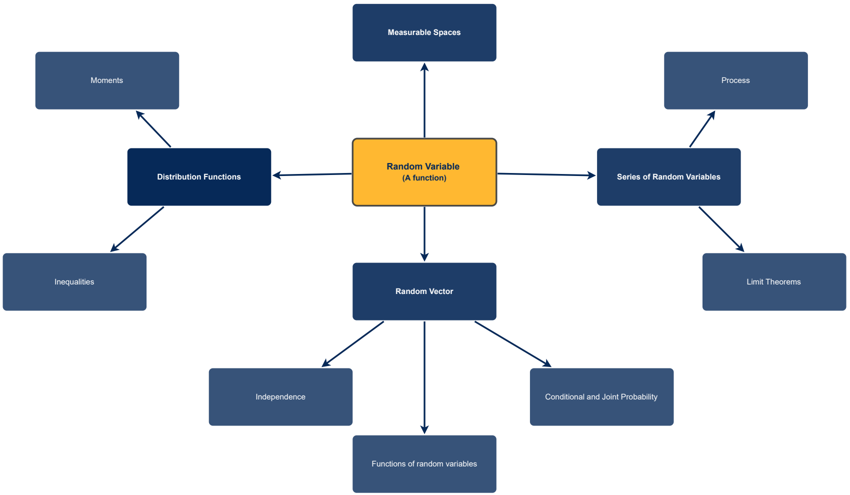


Figure: Roadmap

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

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: $\{H\}$ $\subset \Omega = \{H, T\}$
- Get an even number when rolling a die: $\{2, 4, 6\}$ $\subset \Omega = \{1, 2, \cdots 6\}$

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,

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

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

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

If we imitate discrete probability, then

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

Real numbers

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:

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

countable sum

But an expression such as

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

[0, 1] contains uncountably many values.

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**.

UNIVERSITY OF TORONTO

A Countable Approximation of the Real Line

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

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

Then

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

$\leftarrow$ union of countably many intervals.

These intervals are disjoint, so countable additivity gives

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

countable sum

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,

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

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

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

Therefore,

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

countable sum

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

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

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

countable operation

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

Let $A \in \mathcal{F}$. Since $\mathcal{F}$ is closed under complement, $A^c \in \mathcal{F}$.

Since $\mathcal{F}$ is closed under countable union, $\underbrace{A \cup A^c}_{= \Omega} \in \mathcal{F}$.

Thus $\Omega \in \mathcal{F}$.

Again, taking the complement of Ω, we conclude $\phi = \Omega^c \in \mathcal{F}$.

Prop $\mathcal{F}$ is closed under countable intersection:

If $A_1, A_2, \ldots \in \mathcal{F}$, $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

(pf.) First observe that

$$\left(\bigcap_{i=1}^{\infty} A_i\right)^c = \bigcup_{i=1}^{\infty} A_i^c$$

Since $\mathcal{F}$ is closed under complement, $A_i^c \in \mathcal{F}$ for all i.

Since $\mathcal{F}$ is closed under countable union, $\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F}$.

Since $\mathcal{F}$ is closed under complement,

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c \in \mathcal{F}.$$

//

Probability measures

algebra

Measures on σ-field

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

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

countably disjoint sets

countable additivity

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

Probability measures

Measures on σ-field

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

- $\mu(\varnothing) = 0$,
- If $A_1, A_2, \dots \in \mathcal{F}$ and $A_i \cap A_j = \varnothing$, 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: $A \subseteq B \quad \Rightarrow \quad \mu(A) \leq \mu(B)$
- Subadditivity: $A \subseteq \cup_{i=1}^{\infty} A_i \quad \Rightarrow \quad \mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$ ← do not assume A_i's are disjoint.
- Continuity from below: $A_i \nearrow A \quad \Rightarrow \quad \mu(A_i) \nearrow \mu(A)$
- Continuity from above: $A_i \searrow A$ and $\mu(A_i) < \infty \quad \Rightarrow \quad \mu(A_i) \searrow \mu(A)$

$A_1 \subset A_2 \subset A_3 \subset \cdots$, $\bigcup_{i=1}^{\infty} A_i = A$

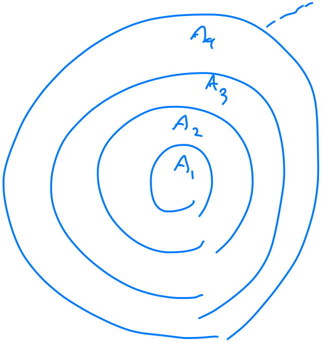
Probability measures

Proof of continuity from below:

Construct disjoint B_i by

$$B_i = A_i \setminus A_{i-1} \quad i \geq 2 \ \checkmark$$

$$B_1 = A_1$$



Since $A = \bigcup_{i \geq 1}^{\infty} B_i$ and by countable additivity

$$\mu(A) = \sum_{i \geq 1}^{\infty} \mu(B_i) = \sum_{i \geq 1}^{\infty} \mu(A_i \setminus A_{i-1})$$

$$= \mu(A_i) - \mu(A_{i-1})$$

$$= \lim_{n \to \infty} \sum_{i \geq 1}^{n} \left(\mu(A_i) - \mu(A_{i-1}) \right) = \lim_{n \to \infty} \mu(A_n) \quad //$$

Probability measures

If $A_i \searrow A$, $\mu(A_i) < \infty$, then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$

Proof of continuity from above:

Define $B_i = A_1 \setminus A_i$, then $B_i \nearrow A_1 \setminus A$. $\longrightarrow$ we reduce to continuity from below.

By continuity from below, we have

$$\lim_{n\to\infty} \mu(B_n) = \mu(A_1 \setminus A)$$

$$\Longleftrightarrow \lim_{n\to\infty} \mu(A_1 \setminus A_n) = \mu(A_1 \setminus A)$$

$$\overset{①}{\Longleftrightarrow} \lim_{n\to\infty} \left(\mu(A_1) - \mu(A_n)\right) = \mu(A_1) - \mu(A)$$

$$\overset{②}{\therefore} \lim_{n\to\infty} \mu(A_n) = \mu(A)$$

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

① $\mu(A \setminus B) = \mu(A) - \mu(B)$ holds if $\mu(A) < \infty$ and $B \subset A$

② $\infty - a = \infty - b$ is possible even if $a \neq b$

Probability measures

Examples:

$\Omega = \{\omega_1, \omega_2, \cdots\}$, $A = \{\omega_{a_1}, \cdots, \omega_{a_i}, \cdots\} \Rightarrow \mu(A) = \sum_{j=1}^{\infty} \mu(\omega_{a_j})$.
Therefore, we only need to define $\mu(\omega_j) = p_j \geq 0$.
If further $\sum_{i=1}^{\infty} p_j = 1$, then μ is a probability measure.

- Toss a coin:

$\Omega = \{H, T\}$, $\mu(H) = p_H$, $\mu(T) = p_T$ when $p_H + p_T = 1$,

Then μ is a probability measure on Ω.

- Roll a die:

$\Omega = \{1, 2, \cdots, 6\}$, $\mu(1) = p_1$, $\cdots$, $\mu(6) = p_6$, $\sum_{i=1}^{6} p_i = 1$

Then μ is a probability measure.

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

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

complement

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:

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

UNIVERSITY OF TORONTO

Generating a σ-Algebra

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

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

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

→ necessary for continuous probability

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

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

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

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

$$[a, b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n}\right)$$

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?
- $P(A \mid B)$ 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?
- $P(A \mid B)$ is the conditional probability of the event A given B.

Example:

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

Conditional probability

Definition

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

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

Conditional probability

Multiplication rule

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

Generalization:

Law of total probability

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

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

Problem Set

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

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.)