Statistical Sciences

DoSS Summer Bootcamp Probability Module 1

Ichiro Hashimoto

University of Toronto

July 8, 2025

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

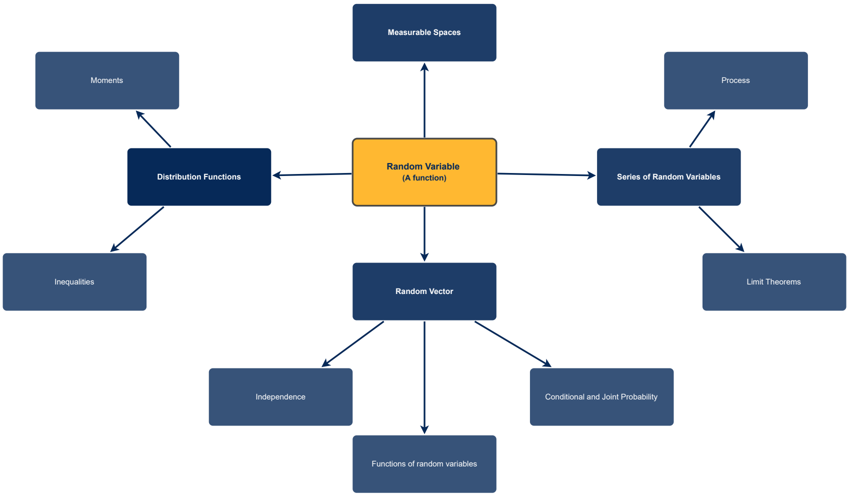


Figure: Roadmap

UNIVERSITY OF TORONTO

Outline

- Measurable spaces
  - Sample Space
  - σ-algebra
- Probability measures
  - Measures on σ-field
  - Basic results

Today

- Conditional probability
  - Bayes' rule
  - Law of total probability

Tomorrow?

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

Event

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

Examples:

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

What is the motivation for developing measure theory?

1) Observation from simple examples

Tossing a coin twice (discrete)

$\Omega = \{HH, HT, TH, TT\}$, $\rightarrow$ discrete.

$\mathbb{P}(HH) = \mathbb{P}(HT) = \mathbb{P}(TH) = \mathbb{P}(TT) = \frac{1}{4}$

Let X = the number of H.

$\mathbb{P}(X=0) = \mathbb{P}(X=2) = \frac{1}{4}$, $\mathbb{P}(X=1) = \frac{1}{2}$

$$\mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2) = 1$$

$$\mathbb{E}X = \frac{1}{4}\cdot 0 + \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 = 1$$

Gaussian (continuous)

Let $X \sim \mathcal{N}(\mu, \sigma^2)$

Density $p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

$$\int_{-\infty}^{\infty} p(x)\,dx = 1$$

$$\mathbb{E}X = \int_{-\infty}^{\infty} x\,p(x)\,dx = \mu$$

Discrete

$$\mathbb{P}(X \le x) = \sum_{\ell \le x} \mathbb{P}(X = \ell)$$

$$\mathbb{E} X = \sum_{k=-\infty}^{\infty} k\, \mathbb{P}(X = k)$$

Continuous

$$\mathbb{P}(X \le x) = \int_{-\infty}^{x} p(x)\, dx$$

$$\mathbb{E} X = \int_{-\infty}^{\infty} x\, p(x)\, dx$$

Q. Is there any way to explain both in a unified manner?

2) Further Observation

If $A \cap B = \phi$, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ --- (*)

For a discrete case, $\{X = k\}$ are disjoint.

Repeating (*),

$$1 = \mathbb{P}(\Omega) = \underbrace{\sum_{k=-\infty}^{\infty} \mathbb{P}(X = k)}_{\text{countable sum}}$$

But for continuous case.

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

Therefore,

uncountable sum cannot be defined well.

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

contradiction?

$\Longrightarrow$ uncountable sum is problematic!

Measurable spaces

σ-algebra

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

(i) • If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, $\longrightarrow$ complement is also in $\mathcal{F}$

(ii) • If $A_1, A_2, \dots \in \mathcal{F}$, then $\underbrace{\cup_{i=1}^{\infty}}_{\text{countable union}} A_i \in \mathcal{F}$. $\longrightarrow$ Countable union is also in $\mathcal{F}$

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

(Pf) Let $A \in \mathcal{F}$.

By (i), $A^c \in \mathcal{F}$.

By (ii), $\Omega = A \cup A^c \in \mathcal{F}$

Again by (i), $\emptyset = \Omega^c \in \mathcal{F}$ //

Construction of Probability Theory

Outline

1) Define the collection of subsets of Ω, $\mathcal{F}$ (σ-algebra) on which we define "probability measure".

2) Define as a function

$$\mathbb{P} : \mathcal{F} \to [0,1]$$

which has "countable additivity".

3) $(\Omega, \mathcal{F}, \mathbb{P})$ is called "probability triple".

Ω: sample space.
$\mathcal{F}$: σ-algebra
$\mathbb{P}$: 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.

contable additivity

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 $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

A_i must be disjoint

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

Properties:

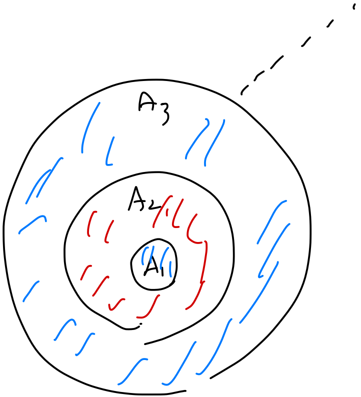
- Monotonicity: $A \subseteq B \quad \Rightarrow \quad \mu(A) \leq \mu(B)$
- Subadditivity: $A \subseteq \cup_{i=1}^{\infty} A_i \quad \Rightarrow \quad \mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$ union bond in probability.
- Continuity from below: $A_i \nearrow A \quad \Rightarrow \quad \mu(A_i) \nearrow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

UNIVERSITY OF TORONTO

Probability measures

Proof of continuity from below: $A_i \nearrow A$

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



Let $B_i = A_i \setminus A_{i-1}$, $i \geq 2$, $B_1 = A_1$.

Then B_i's are disjoint but $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i = A$.

By countable additivity

$$\mu(A) = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i)$$

$$= \sum_{i=2}^{\infty} \left\{ \mu(A_i) - \mu(A_{i-1}) \right\} + \mu(A_1)$$

$$= \cancel{\sum_{i=1}^{\infty}} \mu(A_i)$$

$$\lim_{i \to \infty}$$

$$= \lim_{n \to \infty} \left\{ \sum_{i=2}^{n} \left\{ \mu(A_i) - \mu(A_{i-1}) \right\} + \mu(A_1) \right\}$$

$$= \lim_{n \to \infty} \left\{ \cancel{\mu(A_1)} + \left(\cancel{\mu(A_2)} - \cancel{\mu(A_1)} \right) + \left(\cancel{\mu(A_3)} - \cancel{\mu(A_2)} \right) + \cdots + \left(\underbrace{\mu(A_n)}_{\text{remain}} - \cancel{\mu(A_{n-1})} \right) \right\}$$

$$= \lim_{n \to \infty} \mu(A_n)$$

Probability measures

Proof of continuity from above:

$A_i \searrow A \iff A_1 \supset A_2 \supset A_3 \supset \cdots\cdots, \bigcap_{i=1}^{\infty} A_i = A.$

Let $B_i = A_1 - A_i$, we have $B_i \nearrow A_1 \setminus A$.

By the continuity from below

$$\lim_{i\to\infty} \mu(B_i) = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \mu(A_1 \setminus A).$$

Since $\mu(A_1 \setminus A) = \mu(A_1) - \mu(A)$ and

$$\mu(B_i) = \mu(A_1 \setminus A_i) = \mu(A_1) - \mu(A_i),$$

$$\mu(A_1) - \lim_{i\to\infty} \mu(A_i) = \mu(A_1) - \mu(A)$$

$$\therefore \quad \lim_{i\to\infty} \mu(A_i) = \mu(A).$$

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

Probability measures

countable additivity.

Examples:

discrete case.

$\Omega = \{\omega_1, \omega_2, \cdots\}$, $A = \{\omega_{a_1}, \cdots, \omega_{a_i}, \cdots\} \Rightarrow \mu(A) = \sum_{j=1}^{\infty} \mu(\omega_{a_j})$.
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:

  Define $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$. Then $\mathbb{P}$ is a probability measure.

- Roll a die:

  $\mathbb{P}(1) = \mathbb{P}(2) = \cdots = \mathbb{P}(6) = \frac{1}{6}$

  Then $\mathbb{P}$ is a probability measure.

UNIVERSITY OF TORONTO

Q. How does $(\Omega, \mathcal{F}, \mathbb{P})$ provide a unified theory?

sample space, σ-algebra, probability measure.

Observation

(Discrete r.v. X)

$$1 = \mathbb{P}(\Omega) = \sum_{k=-\infty}^{\infty} \mathbb{P}(X=k)$$

$$\mathbb{E}X = \sum_{k=-\infty}^{\infty} k\, \mathbb{P}(X=k)$$

(Continuous r.v.)

$$1 = \mathbb{P}(\Omega) = \int_{-\infty}^{\infty} p(x)\, dx$$

$$\mathbb{E}X = \int_{-\infty}^{\infty} x\, p(x)\, dx.$$

Consider Approximation

$$\Omega = \bigcup_{i=-\infty}^{\infty} \left\{ X \in \left[\frac{i}{n}, \frac{i+1}{n}\right) \right\}$$ (becomes finer as $n \nearrow \infty$)

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

Approximation of Expectation

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

should become precise as $n \to \infty$

$\Downarrow$

We can instead define $\mathbb{E}X$ as the limit of the approximation above.

(informal measure-theoretic definition of expectation)

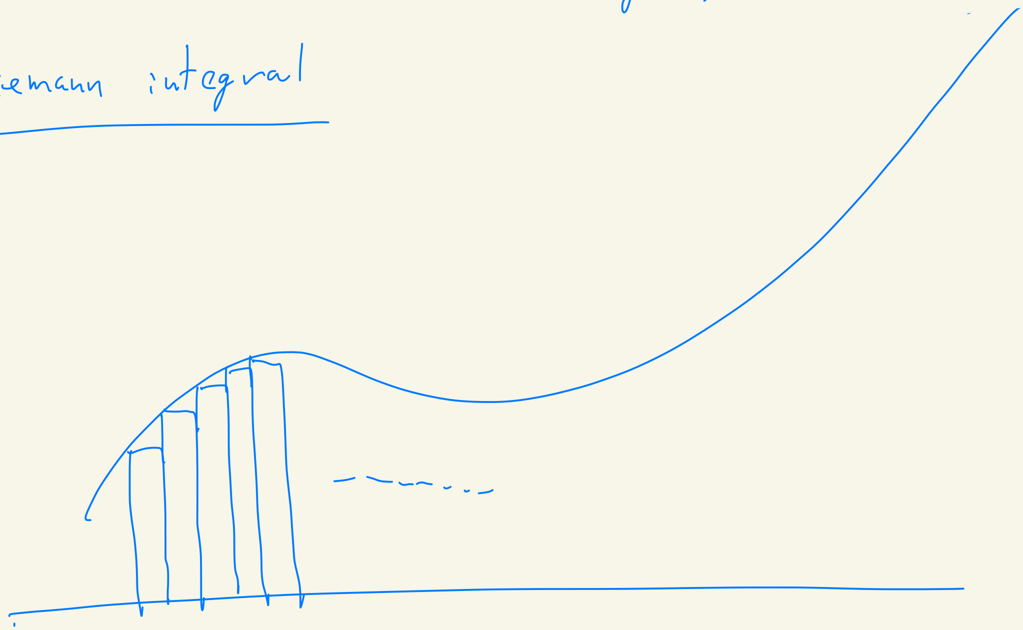
$$\mathbb{E}X \overset{\text{def}}{=} \lim_{n\to\infty} \sum_{i=-\infty}^{\infty} \frac{i}{n}\, \mathbb{P}\left(X \in \left[\frac{i}{n}, \frac{i+1}{n}\right)\right)$$

this looks similar to

Riemann integral

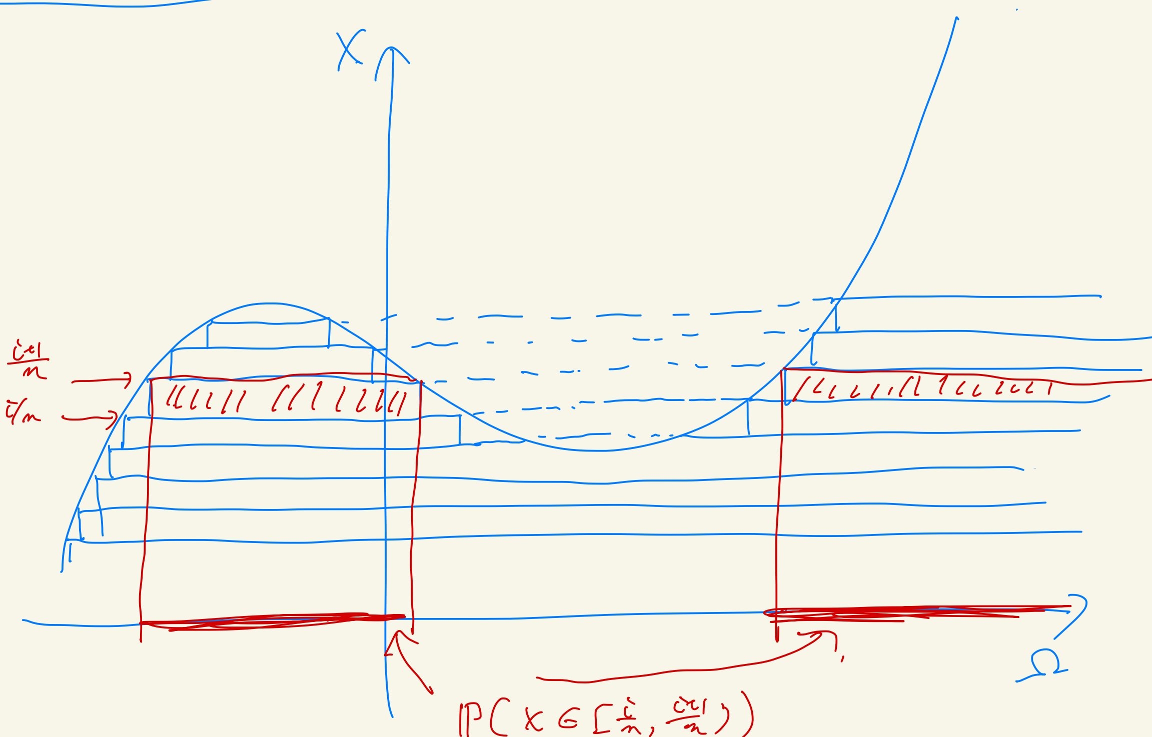
Difference between Riemann integral.

Riemann integral



Measure theory

$X : \Omega \to \mathbb{R}$



$$\mathbb{E}X \overset{\text{def}}{=} \lim_{n\to\infty} \sum_{i=-\infty}^{\infty} \frac{i}{n} \mathbb{P}\left(X \in \left[\frac{i}{n}, \frac{i+1}{n}\right]\right) = \int_{\Omega} x \, d\mathbb{P}$$

We can indeed show

$$\mathbb{E}X = \sum_{i=1}^{\infty} h_i \, \mathbb{P}(X = h_i) \quad \text{in discrete case}$$

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \, p(x) \, dx \quad \text{in cont. case.}$$

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

Bayes' rule

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

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-field $\mathcal{F}$, 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-field.

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.)