

Outline

- 1) Syllabus
- 2) Course goals
- 3) Measure theory basics

Measure theory basics

Measure theory is a rigorous grounding for probability theory [subject of 205A]

Simplifies notation & clarifies concepts, especially around integration & conditioning [Pset 0]

Given a set X , a measure μ maps subsets $A \subseteq X$ to non-negative numbers $\mu(A) \in [0, \infty]$

Example X countable (e.g. $X = \mathbb{Z}$)

Counting measure $\#(A) = \# \text{ points in } A$

Example $X = \mathbb{R}^n$

Lebesgue measure $\lambda(A) = \int_A \dots \int dx_1 \dots dx_n$
 $= \text{Volume}(A)$

Standard Gaussian distribution:

$$\begin{aligned} P_z(A) &= \mathbb{P}(Z \in A) \quad \text{where } Z \sim N(0, 1) \\ &= \int_A \phi(x) dx \quad \phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \end{aligned}$$

NB Because of pathological sets, $\lambda(A)$ can only be defined for certain subsets $A \subseteq \mathbb{R}^n$ [HW 0, Prob. 3]

In general, the domain of a measure μ is a collection of subsets $\mathcal{F} \subseteq 2^X$ (power set)

\mathcal{F} must be a σ -field meaning it satisfies certain closure properties (not important for us)

$$\textcircled{1} X \in \mathcal{F}$$

$$\textcircled{2} \text{ If } A \in \mathcal{F} \text{ then } X \setminus A \in \mathcal{F}$$

$$\textcircled{3} \text{ If } A_1, A_2, \dots \in \mathcal{F} \text{ then } \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$$

$$\underline{\text{Ex}}: X \text{ countable, } \mathcal{F} = 2^X$$

$$\underline{\text{Ex}}: X = \mathbb{R}^n, \mathcal{F} = \text{Borel } \sigma\text{-field } \mathcal{B}$$

$$\mathcal{B} = \text{smallest } \sigma\text{-field including all open rectangles} \\ (a_1, b_1) \times \dots \times (a_n, b_n) \quad a_i < b_i \quad \forall i$$

Given a measurable space (X, \mathcal{F}) a measure is a map $\mu: \mathcal{F} \rightarrow [0, \infty]$ with

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) \quad \text{for disjoint } A_1, A_2, \dots \in \mathcal{F}$$

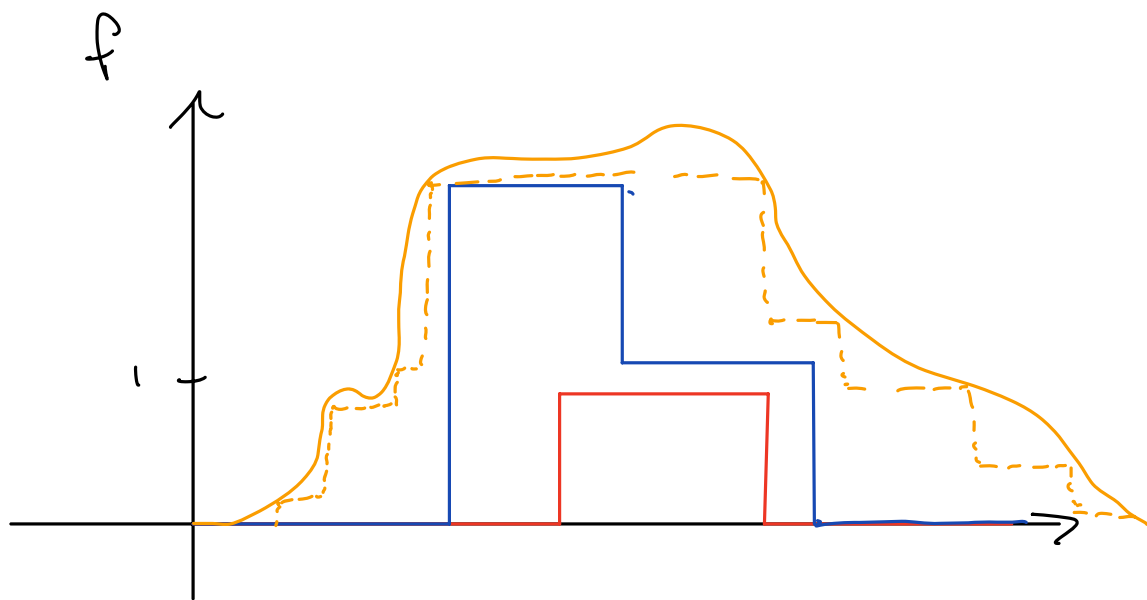
$$\mu(\emptyset) = 0$$

μ probability measure if $\mu(X) = 1$

Integrals

Measures let us define integrals that put weight $\mu(A)$ on $A \subseteq X$

Define $\int 1_{\{x \in A\}} d\mu(x) = \mu(A)$, extend to other functions by linearity & limits:



Indicator $\int 1_{\{x \in A\}} d\mu(x) = \mu(A)$

Simple Function $\int \left(\sum c_i 1_{\{x \in A_i\}} \right) d\mu(x) = \sum c_i \mu(A_i)$

"Nice enough" (measurable) function $\int f(x) d\mu(x)$ approximated by simple functions

Examples:

Counting: $\int f d\# = \sum_{x \in X} f(x)$

↙ Lebesgue
integral

Lebesgue: $\int f d\lambda = \int \dots \int f(x) dx_1 \dots dx_n$

Gaussian: Note $\int 1_A(x) dP_z(x) = P_z(A) = \int_{-\infty}^{\infty} 1_A \phi dx$

By extension,

$$\int f dP_z = \int f(x) \phi(x) dx = \mathbb{E}[f(z)]$$

To evaluate $\int f dP_z$ rewrite as $\int f \phi dx$.
↙ density [can't always
do this]
e.g. Binom

It is nice to turn integrals we care
about into Lebesgue integrals. When
can we do this?

Densities

λ and P above are closely related. Want to make this precise.

Given (X, \mathcal{F}) , two measures P, μ

We say P is absolutely continuous wrt μ
if $P(A) = 0$ whenever $\mu(A) = 0$

Notation: $P \ll \mu$ or we say μ dominates P

If $P \ll \mu$ then (under mild conditions) we can always define a density function

$p: X \rightarrow [0, \infty)$ with

$$P(A) = \int_A p(x) d\mu(x)$$

$$\int f(x) dP(x) = \int f(x) p(x) d\mu(x)$$

Sometimes written $p(x) = \frac{dP}{d\mu}(x)$, called
Radon - Nikodym derivative

Densities are very useful:

Turn $\int f(x) dP(x)$ into something we know how to evaluate, such as

$$1) \int_{\mathcal{X}} f(x) p(x) dx \quad (X \text{ continuous, } \mathcal{X} \subseteq \mathbb{R}^n)$$

$p(x)$ called probability density function (pdf)

$$2) \sum_{x \in \mathcal{X}} f(x) p(x) \quad (X \text{ discrete, } \mathcal{X} \text{ countable})$$

$p(x)$ called probability mass function (pmf)

Often define distributions by giving their density wrt some known measure, e.g.

Ex: Binom (n, θ) pmf: $p(x) = \theta^n (1-\theta)^{n-x} \binom{n}{x}$, $x = 0, \dots, n$

(density p wrt counting measure on $\mathcal{X} = \{0, \dots, n\}$)

Note this dist. has no density wrt Lebesgue:

$$\int_{\{0, \dots, n\}} p(x) dx = 0 \quad \text{for any function } p$$

Probability Space, Random Variables

Typically, we set up a problem with multiple random variables having various relationships to one another.

Want to be able to talk about the "prob. that something happens"

Convenient setup:

R.V.s as functions of an abstract "outcome" ω

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space

$\omega \in \Omega$ called outcome

$A \in \mathcal{F}$ called event

$\mathbb{P}(A)$ called probability of A

A random variable is a function $X: \Omega \rightarrow \mathcal{X}$

We say X has distribution Q ($X \sim Q$)

$$\begin{aligned} \text{if } \mathbb{P}(X \in B) &= \mathbb{P}(\{\omega: X(\omega) \in B\}) \\ &= Q(B) \end{aligned}$$

More generally, could write events involving many R.V.s:

$$\mathbb{P}(X > Y > Z \geq 0) = \mathbb{P}(\{\omega: \dots\})$$

The expectation is an integral w.r.t. \mathbb{P}

$$\mathbb{E}[f(X, Y)] = \int_{\Omega} f(X(\omega), Y(\omega)) d\mathbb{P}(\omega)$$

To do real calculations we must eventually boil
 \mathbb{P} or \mathbb{E} down to concrete integrals/sums/etc.

If $\mathbb{P}(A) = 1$ we say A occurs almost surely

More in Keener ch. 1, much more in Stat 205A