Outline

1) Syllabus
2) Measure theory basics
Measure theory basics

Measure theory is a rigorous grounding for probability theory \([\text{subject of 205A}]\)

Simplifies notation & clarifies concepts, especially around integration & conditioning \([\text{Set 0}]\)

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

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

- Counting measure \(\#(A) = \#\) points in \(A\)

Example \(X = \mathbb{R}^n\)

- Lebesgue measure \(\lambda(A) = \int_A \ldots \int dx_1 \ldots dx_n\)

- Volume \((A)\)

Standard Gaussian distribution:

\[ P(A) = \Pr(Z \in A) \text{ where } Z \sim N(0, 1) \]

\[ = \int_A \phi(x) \, dx \]

\[ \phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \]

\text{NB} Because of pathological sets, \(\lambda(A)\) can only be defined for certain subsets \(A \subseteq \mathbb{R}^n\) \([\text{HWW, Prob. 3}]\)
In general, the domain of a measure $\mu$ is a collection of subsets $\mathcal{F} \subseteq 2^X$ (power set) of $X$ must be a $\sigma$-field meaning it satisfies certain closure properties (not important for us).

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

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

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

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

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

$\mu$ probability measure if $\mu(X) = 1$

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


Define \( \int 1\{x \in A\} \, dm(x) = m(A) \), extend to other functions by linearity & limits:

\[
\int f \, dm(x) = \int \sum c_i 1\{x \in A_i\} \, dm(x) = \sum c_i m(A_i)
\]

“Nice enough” \( f(x) \, dm(x) \) approximated by simple functions

**Examples:**

- **Counting:** \( \int f \, dx = \sum_{x \in x} f(x) \)
- **Lebesgue:** \( \int f \, dx = \ldots \int f(x) \, dx \ldots \)
- **Gaussian:** \( \int f \, \varphi(x) \, dx = \mathbb{E}[f(z)] \)

To evaluate \( \int f \, dp \) rewrite as \( \int f \phi \, dx \).
Densities

\(\lambda\) and \(P\) above are closely related. Want to make this precise.

Given \((X, \mathcal{F})\), two measures \(P, m\)
We say \(P\) is \textit{absolutely continuous} wrt \(m\)
if \(P(A) = 0\) whenever \(m(A) = 0\)

Notation: \(P \ll m\) or we say \(m\) dominates \(P\)

If \(P \ll m\) then (under mild conditions) we can always define a \textit{density function}
\[ p : X \rightarrow [0, \infty) \text{ with} \]
\[ P(A) = \int_A p(x) \, dm(x) \]
\[ \int f(x) \, dP(x) = \int f(x) p(x) \, dm(x) \]

Sometimes written \(p(x) = \frac{dP}{dm}(x)\), called \textit{Radon–Nikodym derivative}

Useful to turn \(\int f(x) \, dP\) into \(\int f(x) p(x) \, dm\) if we know how to calculate integrals \(dm\)

If \(P\) prob., \(m\) Lebesgue:
\(p(x)\) called \textit{prob. density function (pdf)}

If \(P\) prob., \(m\) counting:
\(p(x)\) called \textit{prob. mass function (pmf)}
Probability Space, Random Variables

Typically, we set up a problem with multiple random variables having various relationships to one another, convenient to think of them as functions of an abstract "outcome" $\omega$.

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 $X: \Omega \to \mathcal{X}$

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

$\mathbb{P}(X \in B) = \mathbb{P}(\{ \omega : X(\omega) \in B \}) = Q(B)$

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

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

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 $P$ or $E$ down to concrete integrals/sums/etc.

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

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