Office hours for next week:
Tuesday 3-4 PM
Wednesday 3-4 PM
Permanent office hours TBA

Basics of probability theory:

Probability space
- Set of possible outcomes that are uncertain in advance of the experiment/observation, etc.

Probability measure
- Function on the "nice" subsets of $\Omega$

$\sigma$-algebra: $\mathcal{B}(\Omega)$, what the nice subsets are

For discrete probability spaces, the sigma algebra is often just the space of all subsets of $\Omega$.

$\mathcal{B}(\Omega) = 2^\Omega$

Power set of $\Omega$

And there really are no technicalities.

In setting up a probability model, one specifies the probability space, the sigma-algebra, and the probability measure. This is not a derivation or a calculation, it is defining the model.

However, there are some consistency conditions that must be satisfied by the sigma algebra and probability measure.
The sigma algebra must satisfy the following conditions:

1) \( A \in \mathcal{B}(\mathbb{R}) \Rightarrow A^c \in \mathcal{B}(\mathbb{R}) \) 

2) \( \bigcap_{i \in I} A_i \in \mathcal{B}(\mathbb{R}) \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{B}(\mathbb{R}) \) 

for finite or countable \( I \) (\( A_i \) or \( A_i \) or \( A_i \)).

I countable means \( I \) is isomorphic to some subset of \( \mathbb{N} \).

Or in other words, there is a way to exhaustively enumerate \( I \) by labeling it sequentially with the integers, starting from 0.

Examples of countable sets:

\( \mathbb{N}, \mathbb{Q}, \mathbb{Z} \) but not \( [0,1] \)

\( \emptyset \in \mathcal{B}(\mathbb{R}) \)

Also the probability measure has the following conditions:

1) \( P(A) \geq 0 \) for \( A \in \mathcal{B}(\mathbb{R}) \)

2) \( P(\mathbb{R}) = 1 \)

3) For disjoint countable collection of sets \( \{ A_i \}_{i \in I} \in \mathcal{B}(\mathbb{R}) \),

\( A_i \cap A_{i'} = \emptyset \) for \( i \neq i' \)

\( P(\bigcup_{i \in I} A_i) = \sum_{i \in I} P(A_i) \) (mutually exclusive)

Some elementary consequences of these axioms:

a) \( P(\emptyset) = 0 \)

b) \( P(A^c) = 1 - P(A) \) for \( A \in \mathcal{B}(\mathbb{R}) \)

c) \( P(A \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \)
Generalization to unions of more than 2 sets: inclusion-exclusion formula

\[ P(A_1 \cap A_2) = ?? \]

In general, cannot say anything without further assumption!!

**Independence**

A collection of sets \( \{ A_i \}_{i \in I} \subseteq \mathcal{B}(\Omega) \)

Are said to be independent provided:

\[
P\left( \bigcap_{i \in J} A_i \right) = \prod_{i \in J} P(A_i) \quad \text{for all } J \subseteq I
\]

That's the mathematical definition of independence. Intuitively, we think of events as independent when the realization of one event has nothing to do with the realization of another event.

**Conditional probability:**

For \( A, B \subseteq \mathcal{B}(\Omega) \)

\[
P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{provided } P(B) \neq 0.
\]

Notice that if \( A \) and \( B \) are independent, then

\[
P(A \mid B) = \frac{P(A) P(B)}{P(B)} = P(A)
\]

A brief abstract perspective on stochastic processes:

\[ \Xi(t) : T \rightarrow S \]

How do you define a stochastic process (random function)?
Mathematically? There are several possible approaches.

1. **Intuitive:** Think of \( \{ \Xi(t) \}_{t \in T} \) as a collection of random variables, \( t \in T \) where random variables are
2. Think of $X(t)$ itself as a random object, corresponding to an outcome in some probability space of functions. This is somewhat intuitive but difficult to work with in practice because whenever the time domain $T$ is infinite, the probability space will be uncountable (not discrete) and one needs to use measure-theory technicalities.

3. "Monte Carlo" perspective: View the stochastic process actually as a function jointly on time domain and probability space:

$$X = X(t, \omega): T \times \Omega \rightarrow S$$

It is a dynamical process as a function of $t$ and a random variable as a function of $\omega$.

4. "Weak formulation": Any measurement (functional) of the stochastic process is a random variable (function on probability space), it’s the underlying framework for "stochastic finite element methods."

Let’s begin with probably the simplest, most trivial, stochastic process.

Sequence of independent, identically distributed random variables

$\{X_n\}_{n=0}^{\infty}$ with $X_n \in S$

$$P(\{X_n \in B_n\}, n \in I) = \prod_{n \in I} P\{X_n \in B_n\}$$

independent

Identically distributed:

$$P\{X_n \in B\} = P\{X_n \in B\} \text{ for } n \in I$$

and $B \in \mathcal{B}(S)$

Basic primer on random variables:

A random variable $X$ is a "measurable function" on probability space:
Measurable function means that for any nice subset $B$ of $S$

$$B \in \mathcal{B}(S)$$

We can calculate $\mathbb{P}(\{\omega : X(\omega) \in B\})$ for $B \in \mathcal{B}(S)$, meaning

$$\{\omega : X(\omega) \in B\} \in \mathcal{B}(\Omega)$$

It is often useful to "lift" the probability measure from the underlying probability space to probability measures for the random variables themselves on their state space:

We define the probability measure for a random variable $X$ on its state space $S$ by:

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

for $B \in \mathcal{B}(S)$ (nice subset of $S$).

For discrete state spaces, one simply has:

$$\mathbb{P}_X(\{X\}) = \mathbb{P}(\{X(\omega) = x\})$$

$$\mathbb{P}_X(B) = \sum_{x \in B} \mathbb{P}_X$$

For readings, see the first chapter of most books; I will scan in:

- Karlin and Taylor Sec. 1.1
- Kloeden and Platen Sec. 1.1-1.3