Set Algebra:

- **Union -** A or B; $A \cup B$.
- **Intersection -** A and B; $A \cap B$.
- **Complement -** not A; A^C.
- **Difference -** A but not B; $A \backslash B$
- **Disjoint Events aka. mutually exclusive -** events A and B are disjoint if they don't share any outcomes in common (i.e., A and $B = \varnothing$).
- **Subset -** $A \subseteq B$

Trial - a repetition of a random experiment/process. Trials're independent: none gives information about the others; are stable: reuslts could have appeared in any order.

Outcome - a possible result of a trail.

Sample space - the set of all possible outcomes. Often denoted as S.

Event - a set of outcomes of an experiment (i.e., a subset of the sample space).

Probability - is a long run proportion of an outcome in repeated trials.

- Probabilities act as "targets" of estimation
- Proportions based on data "estimate" probabilities. Would approach probabilities if observe infinite trials.

Formally, $\boxed{A \mapsto \mathbb{P}(A), \mathbb{P}(A) \in [0,1]}$ A probability $\mathbb{P}(\cdot)$ on a sample space S is a function that assigns a snumber between 0 and 1 to all events, A in the sample space (i.e., any possible subset of the sample space) and subject to three requirements (axioms):

1. $\mathbb{P}(S) = 1$: probability of *something* in the sample space happening is 1
2. $\mathbb{P}(A) \geq 0, \forall A$
3. A and $B = \varnothing$ (A, B disjoint) $\Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

More takeaways

- $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
- A, B, C are pairwise disjoint $\Rightarrow \mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \text{ and } B)$

Joint Probability - $\mathbb{P}(A \text{ and } B)$ is the joint probability that events A and B occur.

Conditional Probability - $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)}, \mathbb{P}(B) > 0$ the probability of observing event A if (given that) one has observed B. Bear in mind: $\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$ **Product Rule -** $\mathbb{P}(A \text{ and } B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B)$.

Independent Events - A and B are independent if $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$; one event happening doesn't affect the probability of the other event happening. Can easily deduce that $\mathbb{P}(A|B) = \mathbb{P}(A)$ and $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Independence and Disjointness are **NOT** synonyms.

- Independent $\Rightarrow \mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$
- Disjoint $\Rightarrow \mathbb{P}(A \text{ and } B) = 0$. Disjoint events are extremely dependent: If one event occurs, the other cannot.

Random variable - a numerical function on a sample space with probabilities. (Think as a scoring mechanism.)

- Input: an outcome in the sample space
- Output: a number

Discrete RVs - only countably many values are possible

Continous RVs - can take on uncountably infinitely many values

Probability Distribution Function (PDF) - $\boxed{\mathbf{p}_X(x) = \mathbb{P}(X = x)}$ is the probability that the random variable X takes on the value x. I really hate $\mathbf{p}_X(x)$ this styling, so only $\mathbb{P}(X = x)$ moving forward.

Properties of PDFs - any function that satisfies the following conditions is a probability distribution function of a Discrete random variable:

1. $\mathbb{P}(X = x) \geq 0, \forall x \in \mathbb{R}$ (for any real number)
2. $\mathbb{P}(X = x) > 0$ for values that the random variable X can actually take on
3. $\mathbb{P}(X = x) = 0$ for values that aren't possible for the random variable X
4. $\sum_x \mathbb{P}(X = x) = 1$

Expected Value - $\boxed{\mathbb{E}[X] = \sum_x x \cdot \mathbb{P}(X = x)}$

Variance -

$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sum_x (x - \mathbb{E}[X])^2 \cdot \mathbb{P}(X =$

Standard Deviation -