PROBABILITY - EXPERIMENT, SAMPLE SPACE, EVENTS

EXPERIMENT: Any procedure that can be repeated under the same conditions (theoretically) infinite number of times, and such that its outcomes are well defined. By *well defined* we mean we can describe **all** possible outcomes.

Sample outcome or outcome: Any possible outcome of the experiment.

Sample space: The set of all possible outcomes of an experiment. Usually denoted by S.

Event: A subset of the sample space S. Events are usually denoted by capital letters.

ALGEBRA OF SETS

Operations on events (sets): Union, intersection, complement.

Definition: Let A and B be two events over the sample S.

1. The union of A an B is the event (set) $A \cup B$ which elements belong to A or B or both A and B.

2. The intersection of A an B is the event (set) $A \cap B$ which elements belong to both A and B.

3. The complement of A is the event (set) A^C which elements do not belong to A (of course they belong to S.

NOTE: The empty set is denoted by $\emptyset$.

NOTE: We can extend the definition of a union (or intersection) of two events, to any finite number of events $A_1, A_2, \ldots, A_k$ defined over the sample space S.

NOTE: Two events A and B are called mutually exclusive if their intersection is empty, that is $A \cap B = \emptyset$.

1. The union of $A_1, A_2, \ldots, A_k$ is the event (set) $\bigcup_{i=1}^{k} A_i = A_1 \cup A_2 \cup \ldots \cup A_k$ which elements belong to at least one of the sets $A_1, A_2, \ldots, A_k$.

2. The intersection of $A_1, A_2, \ldots, A_k$ is the event (set) $\bigcap_{i=1}^{k} A_i = A_1 \cap A_2 \cap \ldots \cap A_k$ which elements belong to all the sets $A_1, A_2, \ldots, A_k$.

DeMorgan Laws: Treat complement of a union or intersection.

1. The complement of a union of $A_1, A_2, \ldots, A_k$ is the intersection of the complements $A_1^C, A_2^C, \ldots, A_k^C$, that is $(\bigcup_{i=1}^k A_i)^C = \bigcap_{i=1}^k A_i^C$.

2. The complement of an intersection of $A_1, A_2, \ldots, A_k$ is the union of the complements $A_1^C, A_2^C, \ldots, A_k^C$, that is $(\bigcap_{i=1}^k A_i)^C = \bigcup_{i=1}^k A_i^C$.

MEASURE THEORY & PROBABILITY (661)

The more general foundation of Probability Theory is *Measure Theory*, which also plays a fundamental role in Real and Complex Analysis. We'll visit this perspective on Probability Theory throughout the course whenever it provides useful insights on the underlying theory.

BASIC DEFINITIONS

A **probability space** comprises three parts (S,$\mathcal{F}$,P):

1. S is the **sample space**, the set of all **outcomes**. (Some texts use Ω instead of S). Ex: For a coin toss experiment, $S = \{H, T\}$.

2. $\mathcal{F}$ is the σ-**algebra** associated with S. It is the set of all subsets of S (i.e., the set of all **events**, includes S and $\emptyset$), and is closed under countable unions, countable intersections and complementation. Furthermore, $\mathcal{F}$ satisfies:

   (a) if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$, and

   (b) if A_1, A_2, ... are in $\mathcal{F}$, then their union $\bigcup_i A_i$ is also in $\mathcal{F}$.

   Note: Together these conditions imply closure under countable intersections.

3. P is our probability function $P : \mathcal{F} \to [0, 1]$. It associates each event (i.e., each subset of S included in $\mathcal{F}$) with a number between 0 and 1. Furthermore, we require that

   (a) P is non-negative ($P(A) \geq P(\emptyset) = 0,\ \forall A \in \mathcal{F}$),

   (b) P is countably additive, i.e., for for a countable, disjoint set of events A_1, A_2, ... then $P(\bigcup A_i) = \sum_i P(A_i)$), and

   (c) $P(S) = 1$.

A probability space is a special case of the more general **measureable space** (S, $\mathcal{F}$, μ) which is the conceptual foundation for integration theory taught in standard upper division Real and Complex Analysis courses. Probability functions are a kind of **measure**, and probability spaces are special only in the sense that $\mu(S) = 1$. Measure theory deals with the more general case where $\mu(S) \geq \infty$ (for example, our usual notion of "distance" on

$\mathbb{R}$ implies a standard measure known as Lebesgue Measure, e.g., $\mu([a, b]) = b - a$, where $\mu(S) = \mu(\mathbb{R}) = \infty$). Accordingly, probability spaces are relatively "nice" spaces to work with from an analysis perspective!

NOTE: It's worth noting these deeper connections to Real and Complex Analysis, as you may encounter some of them in future classes. Additionally, it's worth remembering that probability functions can be though of as functions that associate probabilities (values in [0,1]) with *sets of possible outcomes* (i.e., *events*), and that there are a few different important spaces (e.g., S vs $\mathcal{F}$) we need to keep track of as we progress through the course.

PROBABILITY - THE PROBABILITY FUNCTION, KOLMOGOROV'S AXIOMS

The probability function is a function defined on the set of events (subsets of S). To each event A, it assigns a real number which is its **probability** P(A). To define (or characterize) a probability function P, it is necessary and sufficient that it satisfies the following axioms.

1. Probability of any event A over the sample space S is nonegative: $P(A) \geq 0$.
2. Probability of the sample space is 1: $P(S) = 1$.
3. The probability of a union of two mutually exclusive events A and B is the sum of their probabilities: $P(A \cup B) = P(A) + P(B)$ for any mutually exclusive events A and B.
4. The probability of a union of infinitely many pairwise disjoint events, is the sum of their probabilities. That is, if $A_1, A_2, \ldots$ are events over S such that $A_i \cap A_j = \emptyset$ for $i \neq j$, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

NOTE: Axioms 1 -3 are enough for finite sample spaces. Axiom 4 is necessary when the sample space is infinite.