Probability Theory– I

[Listen to uploaded audio files too if needed]

[If we consider ourselves as civilized humans, then we should be grateful to Mathematicians P. Laplace and E. Borel for developing a systematic Mathematical theory of probability, which controls our stock-exchange and food-production prediction to avoid hunger and climate-model to save lives and many more.]

In this study material, we will start developing set algebraic probability theory. We will try to be very specific in presentation for faster and clearer understanding.

First let us list a set of symbols generally used and their meaning:

		interpretation
the whole sample space	Ω	sure event (contains all possible outcomes)
the empty set	$\emptyset$	impossible event (contains no outcomes)
intersection	$E_1 \cap E_2$	"events E_1 and E_2 both occur"
union	$E_1 \cup E_2$	"event E_1 or event E_2 occurs"
complement	$E^c = \Omega \setminus E$	"event E does not occur"
subset	$E_1 \subset E_2$	"occurrence of E_1 implies E_2"

Explanation: The whole sample space Ω is the entire event space encompassing *all* possible outcomes. The *subset* $(E_1, E_2, E_1 \cup E_2$ etc. etc.) of whole sample space is the event space.

Let us summarize the main points of probability model:

A probabilistic model includes

- The sample space Ω of an *experiment*
  - set of all possible *outcomes*
  - *finite* or *infinite*
  - *discrete* or *continuous*

Explanations:

An experiment means doing-something or drawing-something (example: one lottery ticket drawing by you from a bunch of available tickets).

An outcome is what you draw out of an experiment (example: your lottery ticket draw can have only two options: you are a winner or you are a loser).

The finite experiment means you start some where and stop somewhere during the experiment. Example: you draw 3 times lottery tickets from a bunch of available tickets. This is finite experiment.

The infinite experiment means you actually never stop doing experiments. Example: you are tossing a coin for ever and counting what number of heads and tails are appearing and you are never stopping the coin toss.

The discrete events outcome means that your experiment and events are countable sets. Example: your coin toss experiment.

The continuous events outcome means that your experiment and events are uncountable sets. Example: the random variations of amplitude of a radio signal wave.

Now, let us summarize the concepts of event and outcome of an experiment.

- An event A is a set of outcomes
  - a subset of the sample space, $A \subset \Omega$.
  - special events: certain event: $A = \Omega$, null event: $A = \emptyset$

  The *set of events* $\mathcal{F}$ is the set of all possible subsets (events A) of Ω.
- A probability law $P(A)$ that defines the likelihood of an event A.

-- What is probability?

The probability is a function defined as follows:

$$P : (A \subset \Omega) \rightarrow [0,1]$$

Explanation: The probability function assigns a numerical number between 0 and 1 (inclusive) to the elements of subset of experiment outcomes. In other words, probability function maps events into [0,1] interval. The mapping can be discrete or can be continuous. If the mapping is discrete then it is discrete probability and if it is continuous then it is continuous probability.

The probability function defined above must satisfy a set of axioms as given below.

-- The axioms of probability:

A probability measure $P(A)$ must satisfy the following axioms:

1. $P(A) \geq 0$ for every event A
2. $P(\Omega) = 1$
3. If $A_1, A_2, \ldots$ are disjoint events, $A_i \cap A_j = \emptyset$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Explanations:

The first axiom says that a probability function can never maps into negative domain. This is obvious from the definition of probability function.

The second axiom says that total probability of entire or whole event set is 1. This is also obvious because, in any case at least one event will appear from Ω due to experiment.

The third axiom says that if we have two disjoint event space (subsets) then the total probability of union of two event spaces is equal to sum of probabilities of individual event spaces.

-- The probability space

Now, it is time to formulate a probability space. A probability space is an algebraic triplet given below:

$\{\Omega, \mathcal{F}, P(A)\}$

* *Note: an algebraic triplet is not ordered tuple in relational algebra. However, we denote the entire event-space first as an element because we will base sigma-algebra on it, as we will see later.*

-- Measuring discrete probability

The probability measure $P(A)$ can be defined by assigning a probability to each single outcome event $\{s_i\}$ (or *elementary event*) such that

$$P(s_i) \geq 0 \text{ for every } s_i \in \Omega$$
$$\sum_{s_i \in \Omega} P(s_i) = 1$$

– Probability of any event $A = \{s_1, s_2, \ldots, s_k\}$ is

$$P(A) = P(s_1) + P(s_2) + \ldots + P(s_k)$$

– If Ω consists of n equally likely outcomes, then $P(A) = k/n$.

Explanations:

The first law says that if you take one by one each and every discrete events so that entire event space is touched upon, then the sum of total probability is unity at maximum.

The second law says that if you take subset of event space then the total probability of events in that subset is equal to the sum of probabilities of individual events in that subset.

The third law is a special type (NOT general form) of assignment of probability in [0,1].
It says that if each discrete events are equally likely to happen (i.e. no bias to anybody) then: (the probability of an event subset) = (total events in subset)/(total events in entire set).

Important: We can consider third special case if and only if it is discrete probability (i.e. countable set).

-- Measuring continuous probability

For continuous Ω, the probability measure $P(A)$ cannot be defined by assigning a probability to each outcome.

- For any outcome $s \in \Omega$, $P(s) = 0$

  Note: A zero-probability event does not imply that the event cannot occur, rather it occurs *very infrequently*, given that the set of possible outcomes is infinite.

- But we can assign the probability to an *interval*.

  For example, to define the uniform probability measure over $(0, 1)$, assign $P((a, b)) = b - a$ to all intervals with $0 < a, b < 1$.

Explanation: The interval (0,1) is not same as [0,1]. The interval (0,1) is called open interval and [0,1] is closed interval.

So the uniform probability (i.e. no bias to anybody in continuous case) between any two points *a* and *b* in (0,1) is *b* - *a*. Of course, we consider *b* > *a*.

Assignments:

Q1. Algebraically prove that if $P(A \subset \Omega) = 1$ then $P(\Omega \setminus A) = 0$. Consider discrete probability for simplicity.

Q2. Prove that if $A_1 \subset A_2 ..... \subset A_n \subset \Omega$ is a finite sequence such that $P(A_i) > P(A_{i+k>0})$ then P can*not* be a discrete uniform probability.

Q3. Prove that if Ω is continuous and $A_1 \subset \Omega, A_2 \subset \Omega$ such that $P(A_1) > 0, P(A_2) > 0$ then $P(A_1 \cap A_2) \geq 0$ if $A_1 \cap A_2 \neq \phi$.