Probability Theory– II
[Listen to uploaded audio files too if needed]

In this study material, we will start dealing with basic set theoretic axiomatic laws of probability spaces.

-- Set Algebraic Laws of Probability

First let us summarize a set of main laws as given below:

- If $A \subset B$ then $P(A) \leq P(B)$
- Complement
$$P(A^c) = 1 - P(A)$$
- Joint probability
$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$
- Union
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
- Union of event bound
$$P\left(\bigcup_{i=1}^{N} A_i\right) \leq \sum_{i=1}^{N} P(A_i)$$
- Total probability law: Let $S_1, S_2, \ldots$ be events that partition Ω, that is, $S_i \cap S_j = \emptyset$ and $\bigcup_i S_i = \Omega$. Then for any event A
$$P(A) = \sum_i P(A \cap S_i)$$

Explanations:

The first law says that the probability assignments should be in a way such that total probability of a subset cannot exceed the main set. This is obvious because, main set has more elements than the subset. The equality in the relation is due to the fact that some of elements in main set and subset can have zero assignment by probability function making the total probabilities equal.

The second law (complement law) says that as the total probability is 1 so the probability of a set complement of *A* should be 1-(probability of the set *A*).

The third law (joint probability) says that the probability of an event to be in set *A* AND in set *B* (both) is equal to sum of probability of each set individually minus the probability of the event in either A or in B. Why so?? The reason is that intersection is counted twice in calculating $P(A)$ and $P(B)$.

The fourth law (union law) says that the probability of an event to be in set A OR in set B is equal to sum of probability of each set individually minus the probability of the event in both A and in B. Why so?? The reason is same as before : to eliminate twice-counted probabilities in calculating $P(A)$ and $P(B)$.

The fifth law (union bound) is a generalization of union law. The ≤ ordering relation is due to the reduction (minus) element in union law due to the reason as stated before. The equality in the relation is due to the fact that some zero probability assignments to some elements may bring them equal both sides by numerical adjustments. OR, all the sets are mutually disjoint.

IMPORTANT NOTE: In fifth law it is considered that there are atleast two sets are *not* mutually disjoint if we restrict to $<$ (less than) relation omitting = (equality) relation.

The sixth law (partition law) says that if an event space is partitioned into mutually disjoint subsets (for example power set) then the total probability is equal to the sum of probabilities of each partition.

IMPORTANT NOTE: This sixth law is using the following mathematical identity:

Let $S \subset \Omega$ be an event set and $A \subset S$. Then $A \cap S = A$ in Ω.

-- Conditional Probability

First let us summarize the conditional probability laws :

- Conditional probability is the probability of an event A, given *partial information* in the form of an event B. It is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \ , \quad \text{with} \quad P(B) > 0$$

  - Conditional probability $P(.|B)$ can be viewed as a probability law on the new universe B.
  - $P(.|B)$ satisfies all the axioms of probability.

$$\begin{aligned} P(\Omega|B) &= 1 \\ P(A_1 \cup A_2|B) &= P(A_1|B) + P(A_2|B) \quad \text{for } A_1 \cap A_2 = \emptyset \end{aligned}$$

- The conditional probability of A given B – the *a posteriori* probability of A – is related to the unconditional probability of A – the *a priori* probability – as

$$P(A|B) = \frac{P(B|A)}{P(B)} P(A)$$

Explanations:

The first law of conditional probability comes from the logic that:

$P(A \mid B).P(B) = P(A \cap B)$

Why so?? What is the logic here??

It is this: (probability of *A* to happen given that *B* happened already).(probability that *B* has happened) = (probability of happening both *A* and *B* simultaneously).

IMPORTANT NOTE: In the above expression the following condition is ENFORCED: $P(B) > 0$ because $P(A \subset \Omega) \in [0,1]$.

The question is: why $P(\Omega \mid B) = 1$?? This is because, $B \subset \Omega$.

The next question is: why $P(A \mid B).P(B) = P(B \mid A).P(A)$??

The main set algebraic reason is: $A \cap B = B \cap A$ (called as commutativity). So, logically we can say:

(probability of *A* given *B* happened).(probability of *B* happened) =
(probability of *B* given *A* happened).(probability of *A* happened)

-- Chain rule extension

Now, we have sufficient knowledge to extend the above observations to a more generalized form. It is called chain rule. It is given below:

$$P\left(\cap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P\left(A_3|A_1 \cap A_2\right)\ldots P\left(A_n|\cap_{i=1}^{n-1} A_i\right)$$

Note here that *n* is finite.

Assignments:

Q1. Prove the chain rule using set algebra.

Q2. Prove that if $A_1, A_2 \subset \Omega$ and $A_1 \cap A_2 \neq \phi$ such that $P(A_1) < P(A_2 \setminus A_1 \cap A_2)$ then $\exists \varepsilon > 0$ such that $P(A_1 \setminus A_1 \cap A_2) \geq \varepsilon^{-1}.P(A_2)$.

Q3. Prove that if Ω is discrete uniform probability space and $A_1 \subset \Omega, A_2 \subset \Omega$ such that $A_1 \cup A_2 = \Omega$ then $[P(A_1) = P(A_2)] \Leftrightarrow [|A_1| = |A_2|]$.