Dr. Ujjwal Kumar Pahari, Assistant Professor, Department of Mathematics

Netaji Nagar Day College

Topic for – Paper MTMA CC-11 (Probability & Statistics)

Basic Concepts of Probability

The term probability literally means chance and the theory of probability deals with the chances of occurrence of phenomena which are unpredictable and random in character.

Outcome: The result of an experiment is called outcome of that experiment.

Random experiment: An experiment E is called a random experiment (i) if all possible outcomes of E are known in advance (ii) it is impossible to predict which outcome will occur at a particular performance of E (iii) the experiment can be repeated, at least conceptually, under identical conditions for an infinite number of times.

Example: tossing a coin, throwing an unbiased die etc. Here the term 'random' refers to the fact that any particular outcome cannot be predicted from the knowledge of experimental conditions, i.e., the outcome is uncertain.

Trial: Any particular performance of a random experiment is called a trial.

Sample space or Event space: The set of all possible outcomes of a given random experiment is called sample space of the random experiment and is denoted by Ω.

Example for a random experiment: For the experiment of tossing a coin, the event space is $\Omega = \{H, T\}$ where H denotes the outcome 'Head' and T denotes the outcome 'Tail'. The elements of Ω are called sample points or even points.

Discrete sample space: A sample space is called discrete if it contains a countable number of sample points.

Example-1. the sample space associated with the rolling a die, $\Omega = \{1, 2, 3, 4, 5, 6\}$ contains finite number of sample points where 'i' represents the face marked 'i', i=1, 2,3,….,6.

Example-2. The sample space $\Omega = \{S_1, S_2, S_3, \dots\dots\dots\}$ where $S_1 = H,\ S_2 = TH,\ S_3 = TTH, \dots\dots\dots.$ associated with the experiment of throwing a coin until a head appears, contains an infinitely countable number of sample points.

Uncountable Sample space: A sample space will be called uncountable if it contains an uncountable number of sample points. Example- if we observe the atmospheric temperature at a specific place, the sample points are uncountable and are represented by the real line.

Event: An event of a given random experiment can be defined as a subset (not any subset) of the corresponding sample space.

Note: if the sample space Ω is discrete (countable), then any subset of Ω can be an event but this is not true when Ω is uncountable.

Equally likely sample points: Let Ω be a finite sample space. Then its sample points are said to be equally likely if no one outcome is more likely to occur than another. Example- In drawing a card from a well shuffled deck of 52 playing cards, any card may appear. We cannot say that a particular card will appear more times than the others. So, 52 cards are equally likely.

Certain or Sure Event: An event of a given RE E is called certain event if it happens in every performance of E. The sample space Ω is called certain event, because every outcome is an element of Ω.

Impossible Event: An event of a given RE E is called impossible event if it can never happen in any performance of E. The null set ∅ is called the impossible event, because no outcome of the experiment can be an element of ∅. In throwing a die, 'face marked T' is an impossible event.

Simple Event: An event A is called simple event if it contains only one element. i.e A can happen only in one way in every performance of the RE.

Composite or compound Event: An event A is called composite event if it contains more than one element. i.e A can happen more than one way in every performance of the RE. Example- A = {1, 3, 5} is composite and B = {3} is Simple Event.

Mutually Exclusive Events: Two events A and B are said to be mutually exclusive if they can never happen simultaneously in every performance of of E, i.e., $A \cap B = \emptyset$. Example- In the experiment of throwing a die, 'even face' and 'odd face' are mutually exclusive events.

Exhaustive set of Events: A collection $\{A_\alpha : \alpha \epsilon I\}$ of events is said to be exhaustive if $\bigcup_{\alpha \epsilon I} A_\alpha = \Omega$ where I is the Index set and Ω is the Sample space. Example- In throwing a die, if A = {1, 3, 5}, B = {2}, C = {4, 6} then the set {A, B, C}is exhaustive set of events.

Classical definition of Probability: Let the event space Ω of a RE E be finite and all the simple events (i.e., event points) are equally likely. Then the probability of an event A (⊆ Ω) is defined as $P(A) = \frac{m}{n}$ where m = total number of simple events, i.e., total no. of sample points/ event points, i.e., total no. of distinct elements of Ω. n = number of simple events favourable to A, i.e., no. of sample points in A, i.e., no. of distinct elements in A.

Limitations: i) This definition is applicable only when the sample space is finite and all the simple events are equally likely.

ii) The definition presumes that all the simple events are **equally likely,** but the term 'Equally Likely' cannot be explained without the prior idea of probability.

Statistical Regularity: Let a random experiment E be repeated N times under identical conditions in which an event A occurs N(A) times. The number N(A) is called the frequency of the event A and the ratio $f(A) = \frac{N(A)}{N}$ is defined as frequency ratio of A.

It is an experimental fact that, if E is repeated a very large number of times, then the frequency ratio gradually stabilizes more or less constant. The tendency of stability of frequency ratio for large values of N is called statistical regularity.

Frequency definition/ Statistical definition/Empirical definition of Probability:

Let a random experiment E be repeated N times under identical conditions in which an event A occurs N(A) times. The number N(A) is called the frequency of the event A and the ratio $f(A) = \frac{N(A)}{N}$ is defined as frequency ratio of A.

Then, on the basis of statistical regularity we can assume that $\lim_{N\to\infty} \frac{N(A)}{N}$ exists finitely and the value of this limit is called the probability of the event A, i.e. $P(A) = \lim_{N\to\infty} \frac{N(A)}{N}$.

To define analytical concept of Probability we define some terms.

Field (or algebra): Let $\mathscr{A}$ be a non-empty collection of subsets of a set Ω. Then $\mathscr{A}$ is called a Field or an algebra iff $\Omega \in \mathscr{A}$ and $\mathscr{A}$ is closed under completion and finite union.

i.e., $\mathscr{A}$ is called a field or an algebra iff

i) $\Omega \in \mathscr{A}$

ii) $A \in \mathscr{A} \Rightarrow A^c \in \mathscr{A}$

iii) If $A_1, A_2, A_3, \ldots\ldots, A_n \in \mathscr{A}$ then $\bigcup_{i=1}^{n} A_i \in \mathscr{A}$

σ −field (or σ −algebra):): Let $\mathscr{A}$ be a non-empty collection of subsets of a set Ω. Then $\mathscr{A}$ is called a σ −field or an $\sigma - algebra$ iff $\Omega \in \mathscr{A}$ and $\mathscr{A}$ is closed under completion and **Countable** union.

i.e., $\mathscr{A}$ is called a σ −field or a $\sigma -$ algebra iff

i) $\Omega \in \mathscr{A}$

ii) $A \in \mathscr{A} \Rightarrow A^c \in \mathscr{A}$

iii) If $A_1, A_2, A_3, \ldots\ldots, A_n, \ldots, \ldots\ldots \in \mathscr{A}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$

Example: i) The smallest σ −algebra of a fixed set Ω is $\{\emptyset, \ \Omega\}$.

ii) The largest σ −algebra of a fixed set Ω is the collection of all subsets of Ω i.e., power set of Ω denoted by$P(\Omega)$.

iii) Let A be a non-empty proper subset of Ω . The smallest σ −algebra containing A is $\{\emptyset, \ \Omega, \mathrm{A}, A^c\}$.

Note: i) Every algebra (field) contains $\emptyset$ and Ω.

ii) Every σ −algebra is an algebra but the converse is not true in general

iii) every σ −algebra is closed under countable union and intersection.

Borel σ −algebra or borel field and Borel sets:

The smallest σ −algebra of elements of R that contains all type of intervals is called Borel σ-algebra or borel field and is denoted by $\mathscr{B}$**(R)** or $\mathscr{B}$**.** Elements of $\mathscr{B}$ are called Borel sets.

Note: i) All type of intervals of R are Borel sets.

ii) A singleton set {a} is a Borel set, where $a \in R$, since

$\{a\} = \bigcap_{n=1}^{\infty}\left(a - \frac{1}{n}, \ a + \frac{1}{n}\right) \quad or \quad \{a\} = (-\infty, \ x] \cap [x, \ \infty)$ etc.

Event: Let Ω be the Sample space connected to a random experiment E. Let $\mathscr{A}$ be the σ −algebra of subsets of Ω. Then any element of $\mathscr{A}$ is called an event and $\mathscr{A}$ is called the **class of events.**

If Ω is a discrete sample space, then $\mathscr{A} = P(\Omega)$, the power set of Ω. Then, since any element of $\mathscr{A}$ is an event, so any subset of Ω is an event.

Measurable Space: The space Ω of all outcomes of a random experiment together with the specification of the σ-field $\mathscr{A}$ of events is called a *measurable space* and is denoted by pair (Ω, $\mathscr{A}$).

Set function: A set function is a function whose domain is a collection of sets. Let $\mathscr{A}$ be a collection of sets. Then a set function f: $\mathscr{A} \rightarrow R$ is said to be

i) additive or finite additive if $A_1, A_2, A_3, \ldots\ldots, A_n \in \mathscr{A}$ and $A_i \cap A_j = \emptyset,\ i, j \in \{1, 2, \ldots\ldots n\}$ then $f(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} f(A_i)$

ii) countably additive if $A_1, A_2, A_3, \ldots\ldots \mathscr{A}$ and $A_i \cap A_j = \emptyset,\ i, j \in \{1, 2, \ldots\ldots\}$ then $f(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} f(A_i)$

Examples of set function: Length of an interval, Area, volume, Measure, Probability etc.

Axiomatic definition of Probability:

Let Ω be the sample space of a random experiment E and $\mathscr{A}$ be the σ −algebra of subsets of Ω. A real valued Set function P: $\mathscr{A} \rightarrow R$ is said to be Probability if

i) $P(A) \geq 0,\ \ \forall\ A \in \mathscr{A}$

ii) $P(\Omega) = 1$

iii) If $A_1, A_2, A_3, \ldots\ldots, A_n, \ldots \ldots$ are countable number of pairwise mutually exclusive events (i.e., $A_i \cap A_j = \emptyset,$ for $i \neq j,\ i, j \in \{1, 2, \ldots\ldots\}$ and $A_i, A_j \in \mathscr{A}$), then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Note: 1. Whenever $A \in \mathscr{A}$, then A is said to be an event.

2. P(A) is said to be the Probability of the event A.

3. The Probability function P is countably additive on $\mathscr{A}$ as seen from axiom (iii).

4. The tuple $(\Omega, \mathscr{A}, P)$ is called **Probability space.**

5. Kolmogorov (1933) gave the axiomatic definition of Probability.

Thus, P is a non-negative real valued set function defined on the σ −algebra of subsets of the sample space Ω and countably additive where $P(\Omega) = 1$.

Some properties of probability:

1. $0 \leq P(A) \leq 1, \quad \forall\ A \in \mathscr{A}$
2. $P(\emptyset) = 0.$
3. $P(A^c) = 1 - P(A).$
4. $For\ any\ two\ events\ \ A\ and\ B \in \mathscr{A}, \quad P(A \cup B) = P(A) + P(B) - P(A \cap B).$
5. $If\ A \subset B,\ then\ P(A) \leq P(B).$
6. If $A_1,\ A_2, A_3, \ldots\ldots, A_n \in \mathscr{A}$, then $P(A_1 \cup A_2 \cup \ \ldots\ldots \cup A_n) \leq P(A_1) + P(A_{2)+}\ \ldots\ldots + P(A_n)$

Note:

i) Probability that **'exactly one of the events A and B will occur'** is:

$$P(A\bar{B} + \bar{A}B) = P(A) + P(B) - 2P(A \cap B)$$

ii) Probability that **'At least one of the events A and B will occur'** is:

$$P(A + B) = P(A) + P(B) - P(A \cap B)$$

iii) If $P(A) = 0$, then we cannot say that A is an impossible event, (in this case, A is said to be **stochastically impossible event**) but if A is an impossible event, then $P(A) = 0$.

Conditional Probability:

Let A and B be any two events connected to a random experiment E. The conditional probability of A given B, denoted by P(A|B) is defined as

$P(A|B) = \frac{P(AB)}{P(B)}$, provided $P(B) \neq 0$. Similarly, we define conditional probability of B given A as $P(B|A) = \frac{P(AB)}{P(A)}$, provided $P(A) \neq 0$.

Hence, if $P(A) \neq 0$ and $P(B) \neq 0$, we have

P(AB) = P(A)P(B|A) = P(B)P(A|B)

Note: Conditional Probability satisfies all the axioms of Probability.

Independence of events: Two events A and B are said to be stochastically independent if $P(A \cap B) = P(A)P(B)$, i.e., if the occurrence of the event A does not affect the probability of occurrence of the event B.

General Multiplication Rule:

If $A_1, A_2, A_3, \ldots\ldots, A_n$ $(n \geq 2)$ be *n* events connected to a Random experiment E, then

$P(A_1, A_2, A_3, \ldots\ldots, A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \ldots \ldots P(A_n|A_1A_2 \ldots \ldots A_n)$, provided the conditional probabilities are defined.

Bayes' theorem

For a given probability space $(\Omega, \mathscr{A}, P)$, if $A_1, A_2, A_3, \ldots\ldots, A_n$ are *n* pairwise mutually exclusive and exhaustive set of events (i.e., $A_i \cap A_j = \emptyset$, for $i \neq j$, $i, j \in \{1, 2, \ldots\ldots n\}$ and $\cup_{i=1}^{n} A_i = \Omega$) where $P(A_i) \neq 0$, then for any event $X \in \mathscr{A}$ for which $P(X) \neq 0$,

$$P(A_i|X) = \frac{P(A_i)P(X|A_i)}{\sum_{r=1}^{n} P(A_r)P(X|A_r)}, \quad i = 1, 2, 3, \ldots \ldots, n.$$

Difference between Bays' theorem and conditional Probability:

Bayes' theorem is useful for those experiments consisting of two stages:

In the **first stage**, the event A_r is assumed to happen for *r*=1, 2, 3, … …, *n* and in the **second stage,** the event *X* is defined in terms of the whole experiment.

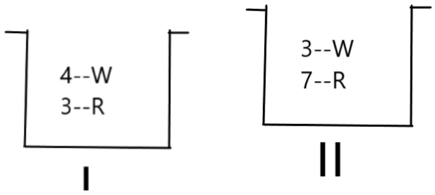
In Bayes' theorem $P(A_i|X)$ is in a backward sense, because we find the Probability of an event A_i defined in terms of the first stage of experiment on the basis of what happens in a later stage of the experiment i.e., on *X*.

But the conditional probability $P(X|A_r)$ is in a forward sense, because here we find the probability of an event *X* defined in the second stage of the experiment on the basis of what happens in the first stage of experiment (i.e., on A_r , *r* =1, 2, 3, … …, *n*) and it is the natural conditioning.

Note: In Bays' theorem, $P(A_i|X)$ is expressed in terms of the natural conditioning given by $P(X|A_r)$ and $P(A_r)$, *r* =1, 2, 3, … …, *n*.

Example: There are two identical urns containing 4 white, 3red balls and 3 white, 7 red balls respectively. An urn is chosen at random and a ball is drawn from it. If the ball drawn is white, what is the probability that it was drawn from the first urn.

Solution: Let A_1, A_2 and X be the events of choosing first urn, second urn and 'the ball drawn is white' respectively.



$\therefore P(A_1) = P(A_2) = \frac{1}{2}$, since the urns are identical.

Here, at the first stage of experiment the events A_1, A_2 are defined and in the second stage, the event *X* is defined. The conditional probabilities $P(X|A_1)$ and $P(X|A_2)$ are in the forward sense as they are natural conditioning.

$$P(X|A_1) = \frac{4}{7}$$

$$P(X|A_2) = \frac{3}{10}$$

In the Bayes' theorem we find $P(A_1|X)$ which is in the backward sense.

$$P(A_1|X) = \frac{P(A_1)P(X|A_1)}{P(A_1)P(X|A_1) + P(A_2)P(X|A_2)}$$

$$= \frac{\frac{1}{2}.\frac{4}{7}}{\frac{1}{2}.\frac{4}{7}+\frac{1}{2}.\frac{3}{10}} = \frac{40}{61}$$

Note: Here the events are mutually exclusive and exhaustive.