APPLICATION OF PROBABILISTIC THEORY

TIMELINE

01

02

03

04

WEEK 1

In week 1 we covered basics of probability, conditional probability,bayes' theorem, independent events, random variables, induced probability space and distribution function along with its properties.

WEEK 2

In week 2 we covered discrete random variables, continuous and absolutely continuous random variables and expectation of a random variable.

WEEK 3

In week 3 we covered moment generating function and some inequalities. Then we were given an assignment based on what all was taught till now.

WEEK 4

In week 4 we covered applications of theories we studied so far,how to use them in real life scenarios and code them. Then we were given as an assignment to research about more applications of these theories.

WEEK 1

We have started developing theory in problem solving perspective,
with not much of proof systems.

Probability theory is useful in studying randomness.
Some definitions:

- **Random Experiment(ε):** All outcomes are known and outcome of a particular trial cannot be predicted , also the experiment can be repeated under identical conditions.
- **Sample space(Ω):** Collection of all possible outcomes of an experiment.
- **Event(E):** Subset of a sample space. E.g. $E \subseteq \Omega$
    - **Class of sets:** A set of sets will be called a class of sets.
    - **Power Set(P(Ω)):** Class of all subsets of a given set (includes Φ)
    - **Set function:** An extended real valued function whose domain is a class of sets.
    - **Countable Set:** A set that has finite number of elements.

Definition 6: Given a random experiment $\mathcal{E}$ and the associated sample space Ω, a probability function P is a set function defined on event space $\mathcal{P}(\Omega)$ satisfying the following three axioms:

Axiom 1: $P(E) \geq 0, \forall E \in \mathcal{P}(\Omega)$;

Axiom 2: if $\{E_i : i \in S\}$ is a countable collection of disjoint events, then

$$P\left(\bigcup_{i \in S} E_i\right) = \sum_{i \in S} P(E_i);$$

Axiom 3: $P(\Omega) = 1$.

The triplet $(\Omega, \mathcal{P}(\Omega), P)$ is called a probability space.

Some results :

Result 1: For any event E

$$0 \leq P(E) \leq 1 \text{ and } P(E^c) = 1 - P(E).$$

Result 2: Let E and F be events such that $E \subseteq F$. Then

$$P(E) \leq P(F) \quad \text{(monotonicity of probability function)}$$

and

$$P(F - E) = P(F) - P(E).$$

Result 3: Let E_1 and E_2 be two events (not necessarily disjoint). Then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2).$$

CONTINUITY OF PROBABILITY FUNCTION

Definition 1: Let $\{E_n\}_{n\geq 1}$ be a sequence of events. The sequence $\{E_n\}_{n\geq 1}$ is said to be

(a) increasing (written as $E_n \uparrow$) if $E_n \subseteq E_{n+1},\ n = 1, 2 \ldots$. In that case the limit of sequence $\{E_n\}_{n\geq 1}$ is defined as $\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n$;

(b) decreasing (written as $E_n \downarrow$) if $E_{n+1} \subseteq E_n,\ n = 1, 2 \ldots$. In that case the limit of sequence $\{E_n\}_{n\geq 1}$ is defined as $\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n$;

(c) monotone if either $E_n \uparrow$ or $E_n \downarrow$.

Some results :

Result 1: Let $\{E_n\}_{n\geq 1}$ be a monotone sequence of events. Then

$$P\left(\lim_{n\to\infty} E_n\right) = \lim_{n\to\infty} P(E_n), \quad \text{(Continuity of probability function)}$$

i.e.,

$$\lim_{n\to\infty} P(E_n) = \begin{cases} P\left(\bigcup_{n=1}^{\infty} E_n\right), & \text{if } E_n \uparrow \\ P\left(\bigcap_{n=1}^{\infty} E_n\right), & \text{if } E_n \downarrow \end{cases}.$$

Result 2: Let $\{E_n\}_{n\geq 1}$ be a sequence of events. Then

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} P(E_n). \quad \text{(generalized Boole's Inequality)}$$

Conditional Probability:

Definition 1: Let A and B be two events. The conditional probability of B given A is defined by

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

provided $P(A) > 0$.

Some results:

Result 1: Let $A \in \mathcal{P}(\Omega)$ be a fixed event such that $P(A) > 0$. Then $P(\cdot|A)$ is a probability function for sample space Ω. It is also a probability function for sample space A.

Result 2 (Theorem of Total Probability): Let $\{E_i : i \in S\}$ be a countable collection of mutually exclusive and exhaustive events ($E_i \cap E_j = \phi$, $\forall\ i \neq j$ and $P(\bigcup_{i \in S} E_i) = 1$). Then, for any other event E,

$$P(E) = \sum_{i \in S} P(E \cap E_i) = \sum_{i \in S} P(E|E_i)P(E_i).$$

Result 3 (Bayes' Theorem):. Let $\{E_i : i \in S\}$ be a collection of mutually exclusive and exhaustive events such that $P(E_i) > 0,\ \forall\, i \in S$. Suppose that E is any other event. Then, for any fixed $j \in S$,

$$P(E_j|E) = \frac{P(E|E_j)\,P(E_j)}{\sum_{i\in S} P(E|E_i)\,P(E_i)}.$$

Remarks:

Remark 1:

- We have
$$P(E_1 \cap E_2) = P(E_1)\,P(E_2|E_1),$$
provided $P(E_1) > 0$;

-
$$\begin{aligned} P(E_1 \cap E_2 \cap E_3) &= P((E_1 \cap E_2) \cap E_3) \\ &= P(E_1 \cap E_2)\,P(E_3|E_1 \cap E_2) \\ &= P(E_1)P(E_2|E_1)P(E_3|E_1 \cap E_2), \end{aligned}$$
provided $P(E_1 \cap E_2) > 0$ (which also guarantees that $P(E_1) > 0$).

Independent Events:

Definition 1: Events E_1, E_2, ..., E_n are said to be

(a) pairwise indenpendent if

$$P\left(E_i \bigcap E_j\right) = P(E_i)P(E_j), \forall i \neq j;$$

(b) mutually independent if $\forall\ k \in \{2, 3, \ldots, n\}$ and distinct $d_1, d_2, \ldots,\ d_k \in \{1, 2, \ldots, n\}$

$$P\left(E_{d_1} \bigcap E_{d_2} \bigcap \cdots \bigcap E_{d_k}\right) = P\left(E_{d_1}\right)\left(E_{d_2}\right)\cdots\left(E_{d_k}\right)$$

Remarks:

When we say that two random experiments are performed independently what it means is that associated events are independent

Suppose that $P\left(E_1\right) > 0$. Then E_1 and E_2 are independent if, and only if,

$$\begin{aligned} P\left(E_1 \bigcap E_2\right) &= P\left(E_1\right) P\left(E_2\right) \\ \Leftrightarrow \qquad \frac{P(E_1 \bigcap E_2)}{P(E_1)} &= P(E_2) \\ \Leftrightarrow \qquad P\left(E_2 | E_1\right) &= P\left(E_2\right) \end{aligned}$$

$\Leftrightarrow$ Conditional probability of E_2 given E_1 is the same as unconditional probability of E_2.

Random Variable:

- $\mathcal{E}$: given random experiment;
- $(\Omega, \mathcal{P}(\Omega), P)$: probability space associated with $\mathcal{E}$;

Example 1:

- $\mathcal{E}$: Tossing a fair can three times independently;
- Sample space $\Omega = \{(\omega_1, \omega_2, \omega_3) : \omega_i \in \{H, T\}, i = 1, 2, 3\}$; here, in $(\omega_1, \omega_2, \omega_3)$, ω_i $(i = 1, 2, 3)$ indicates the outcome of i^{th} toss. Clearly the sample space has $2^3 = 8$ elements;
- Suppose we are interested in number of heads obtained in three tosses, i.e., we are interested the function $X : \Omega \to \mathbb{R}$, where

$$X(w_1, w_2, w_3) = \begin{cases} 0, & \text{if } (\omega_1, \omega_2, \omega_3) = (T, T, T) \\ 1, & \text{if } (\omega_1, \omega_2, \omega_3) \in \{(H, T, T), (T, H, T), (T, T, H)\} \\ 2, & \text{if } (\omega_1, \omega_2, \omega_3) \in \{(H, H, T), (H, T, H), (T, H, H)\} \\ 3, & \text{if } (\omega_1, \omega_2, \omega_3) = (H, H, H) \end{cases}$$

Definition 1: A real valued function $X : \Omega \to \mathbb{R}$ is called a random variable (r.v.).

Distribution Function:

Definition 1: The distribution function (or cumulative distribution function) of a random variable X is a function $F_X\colon \mathbb{R} \to \mathbb{R}$, defined by

$$\begin{aligned} F_X(x) &= P_X((-\infty, x]) \\ &= P(\{\omega \in \Omega : X(\omega) \leq x\}),\ x \in \mathbb{R}. \end{aligned}$$

(a) $0 \leq F_X(x) \leq 1,\ \forall\ x \in \mathbb{R}$;

Result 1: Let $F_X(\cdot)$ be the d.f. of a r.v. X. Then

(a) $F_X(\cdot)$ is non-decreasing;

(b) $F_X(\cdot)$ is right continuous, i.e., $\lim_{n\to\infty} F_X(x + \frac{1}{n})$ exists and

$$\lim_{n\to\infty} F_X(x + \frac{1}{n}) \doteqdot F_X(x+) = F_X(x), \quad \forall\ x \in \mathbb{R};$$

(c) $F_X(-\infty) \doteqdot \lim_{n\to\infty} F_X(-n) = 0$ and $F_X(+\infty) \doteqdot \lim_{n\to\infty} F_X(n) = 1$.

WEEK 2

Discrete Random Variable:

Definition 1: The r.v. X is said to be discrete if there exists a countable set S_X such that

(i)

$$P(\{X = x\}) > 0,\ \forall\, x \in S_X;$$

(ii)

$$P(\{X \in S_X\}) = \sum_{x \in S_X} P(\{X = x\}) = 1.$$

The Set S_X is called the support of random variable X.

A r.v. X is discrete iff

$$\begin{aligned} & P(X \in D_X) && = 1 \\ \Leftrightarrow \quad & \textstyle\sum_{x \in D_X} P(\{X = x\}) && = 1 \\ \Leftrightarrow \quad & \textstyle\sum_{x \in D_X} (F_X(x) - F_X(x-)) && = 1 \\ \Leftrightarrow \quad & \text{sum of sizes of jumps} && = 1. \end{aligned}$$

Probability Mass Function(p.m.f):

Definition 2: Let X be a discrete r.v. with d.f. $F_X(\cdot)$ and support S_X (so that $S_X = D_X,\ P(\{X = x\}) > 0,\ \forall\, x \in S_X$ and $\sum_{x \in S_X} P((\{X = x\}) = 1$. Define $f_X : \mathbb{R} \to \mathbb{R}$ as

$$f_X(x) = \begin{cases} P(\{X = x\}), & \text{if } x \in S_X \\ 0, & \text{otherwise} \end{cases}.$$

The function $f_X(\cdot)$ is called the probability mass function (p.m.f) of r.v. X.

Interdependence of d.f. and p.m.f:

For a discrete r.v. with support S_X and p.m.f $f_X(\cdot)$

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(\{X \le x, X \in S_X\}) \\ &= \sum_{\substack{t \le x \\ t \in S_X}} f_X(t) \end{aligned}$$

$$\begin{aligned} f_X(x) &= P(\{X = x\}) \\ &= \begin{cases} F_X(x) - F_X(x-), & \text{if } x \in S_X \\ 0, & \text{otherwise} \end{cases}, \end{aligned}$$

i.e., p.m.f. $f_X(\cdot)$ determines d.f. $F_X(\cdot)$ and conversely. Thus the probability function $P_X(\cdot)$ induced by r.v. X can be studied through p.m.f $f_X(\cdot)$.

Continuous and Absolutely Continuous Random Variables:

Definition 1 :

(a) The r.v. X is said to be continuous if d.f. $F_X(\cdot)$ is continuous on $\mathbb{R}$;

(b) The r.v. X is said to be absolutely continuous (A.C.) if there exists a non-negative integrable function $f_X : \mathbb{R} \to [0, \infty)$ such that

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt,\ x \in \mathbb{R}.$$

The function $f_X(\cdot)$ is called a probability density function (p.d.f.) of X and the set

$$\begin{aligned} S_X &= \{x \in \mathbb{R} : F_X(x+\epsilon) - F_X(x-\epsilon) > 0, \\ &\qquad \forall\ \epsilon > 0\} \\ &= \{x \in \mathbb{R} : P(\{x - \epsilon < X \le x + \epsilon\}) > 0, \\ &\qquad \forall\ \epsilon > 0\} \end{aligned}$$

is called the support of X.

Expectation of a Random Variable:

- X: a given r.v. with d.f. $F_X(\cdot)$ and p.m.f./p.d.f. $f_X(\cdot)$.

Definition 1:

(a) Let X be a discrete r.v. with support S_X and p.m.f. $f_X(\cdot)$. We say that the expected value of X (denoted by $E(X)$) exists and equals

$$E(X) = \sum_{x \in S_X} x \; f_X(x),$$

provided $\sum_{x \in S_X} |x| f_X(x) < \infty$.

(b) Let X be an A.C. r.v. with p.d.f. $f_X(\cdot)$. We say that the expected value of X (denoted by $E(X)$) exists and equals

$$E(X) = \int_{-\infty}^{\infty} x \; f_X(x) dx,$$

provided $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$.

Some Special Expectations:

- X a r.v.;
- $g : \mathbb{R} \to \mathbb{R}$: a given function;
- Then $Y = g(X)$ is a r.v. and $E(g(X))$ = expected value of $g(X)$.

Some special expectations are:

(i) $\mu_1' = \mu = E(X)$ = mean of (distribution of) X;

(ii) For $r \in \{1, 2, ...\}$, $\mu_r' = E(X^r)$ = r-th moment of X about origin;

(iii) For $r \in \{1, 2, ...\}$, $E(|X|^r)$ = r-th absolute moment of X about origin;

(iv) For $r \in \{1, 2, ...\}$, $\mu_r = E((X - \mu)^r)$ = r-th moment of X about its mean (or r-th central moment);

(v) $\mu_2 = \sigma^2 = E((X - \mu)^2)$ = Variance of X (written as $\mathrm{Var}(x)$).

WEEK 3

Moment Generating Function:

- X: a discrete or A.C. r.v. with p.m.f./p.d.f $f_X(\cdot)$;
- Define
$$A_X = \left\{ t \in \mathbb{R} : E\left(e^{tX}\right) < \infty \right\}.$$
- Clearly $0 \in A_x$, and thus $A_x \neq \phi$.

Definition 1: The moment generating function (m.g.f.) of r.v. X is defined by
$$M_X(t) = E\left(e^{tX}\right), \quad t \in A_X.$$

Some Remarks:

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu'_r .$$

Under the notation and assumptions of above result, let $\psi_X(t) = \ln M_X(t), \; t \in (-h, h)$. Then

$$\begin{aligned} \mu'_1 &= E(X) = \psi_X^{(1)}(0); \\ \text{and} \quad \mu_2 &= \mathrm{Var}(X) = \psi_X^{(2)}(0). \end{aligned}$$

(a) The function $\psi_X(t), \; t \in (-h, h)$, is called the cumulant generating function of X;

(b) For $r = 1, 2, \ldots$

$$\mu'_r = \text{coefficient of } \frac{t^r}{r!} \text{ in Maclaurin's series expansion of } M_X(t).$$

Inequalities:

After this, we studied some inequalities. Some of them are-

Let $g : \mathbb{R} \to [0, \infty)$ be a non-negative function such that $E(g(X)) < \infty$. Then, for any $c > 0$,

$$P(\{g(X) > c\}) \leq \frac{E(g(X))}{c}.$$

(Markov Inequality): Suppose that $E(|X|) < \infty$. Then

$$P(\{|X| > c\}) \leq \frac{E(|X|)}{c}.$$

Result 2 (Chebyshev Inequality): Let X be a r.v. with finite mean $\mu = E(X)$ and finite variance $\sigma^2 = E((X - \mu)^2)$. Then for any $\epsilon > 0$

$$P\left(\{|X - \mu| > \epsilon\sigma\}\right) \leq \frac{1}{\epsilon^2}$$

Definition 1: Let $-\infty \leq a < b \leq \infty$. A function $\phi : (a, b) \rightarrow \mathbb{R}$ is said to be convex (concave) on (a,b) if

$$\phi(\alpha x + (1-\alpha)y) \ \leq \ (\geq) \ \alpha\phi(x) + (1-\alpha)\phi(y), \ \forall \ x, y \in (a, b), \ \alpha \in (0, 1).$$

The function ϕ is said to be strictly convex (strictly concave) if the above inequality is strict.

Result 4 (Jensen Inequality):

Let X be a r.v. with support $S_X \subseteq (a, b)$ and let $\phi : (a, b) \rightarrow \mathbb{R}$ be a convex (concave) function; here $-\infty \leq a < b \leq \infty$. Then

$$E(\phi(X)) \ \geq \ (\leq) \ \phi(E(X)),$$

provided the expectations exist.

Random Vector:

We also briefly introduced with the topic of random vectors and related concept without details. The definition is as follows.

Definition 1.266 (Random Vector)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A function $X = (X_1, X_2, \cdots, X_p) : \Omega \to \mathbb{R}^p$ is called a p-dimensional random vector (or simply, a random vector, if the dimension p is clear from the context). Here, the component functions are denoted by $X_1, X_2, \cdots, X_p$ and each of these are real valued functions defined on the sample space Ω and hence are RVs.

Note 1.267. A 1-dimensional random vector, by definition, is exactly an RV. A p-dimensional random vector is made up of p components, each of which are RVs. Keeping this connection in mind, we repeat the steps of our analysis as done for RVs.

WEEK 4

In this week we had took up concept of various distribution functions. The idea was to study based on application basis. We covered 4 different applications :

- Estimating bounds for loan defaulters with 95% confidence using Binomial Distribution
- Social network analysis using exponential and power law distributions
- Filtering spam mail using Bernoulli Distribution
- Stock market analysis using Normal Distribution and Monte Carlo Simulations

Lets consider the spam mail example,

```
# Sample data (replace with actual word frequencies)
spam_words = {"free": 0.8, "urgent": 0.7, "win": 0.6}

# Function to calculate spam probability based on word occurrence
def is_spam(email_text):
  spam_prob = 1.0  # Initial probability (assuming not spam)
  for word in spam_words:
    if word in email_text:
      spam_prob *= spam_words[word]  # Update probability with word frequency
  return spam_prob

# Example usage
email_text = "This is a legitimate email about your account."
spam_score = is_spam(email_text)

if spam_score > 0.5:  # Threshold for spam classification
  print("This email is likely spam!")
else:
  print("This email seems legitimate.")
```

Assignments

Assignment 1

Assignment 1 included basic questions of all the concepts covered till week 3.

Assignment 2

In this assignment, we were asked to explore out more applications of distributions on grounds of what we covered in Week 4.

PLANS AHEAD

We have more or less stick to what we discussed in our initial presentation. Though we are bit lagging in content covered due to some constraints, we will cover that itself.

Weekly Plan

Week 5

We will dedicate this week to develop some important aspects of Statistics required for further discussion. This will also involve discussing some standard problems of probability and statistics (interview based). We will also introduce some standard models derived from the theory, like in previous week

Week 6

In the 6th week, we delve into Game Theory. Mentees will be introduced to Game Theory's basic concepts and strategies, such as best response, nash equilibrium, mixed strategy equilibrium, etc. We'll also look into popular games like prisoner's dilemma, battle of the sexes, stag-hunt games, etc.

Week 7

We will dedicate the 7th week to Auction Theory. We'll look into different types of Auctions like first-price auctions, second-price auctions, all-pay auctions, etc, and try to devise an optimal strategy for betting to gain maximum profits.

Week 8

This week, we finally go into the real world and apply the concepts learned so far to understand and build models for a real-life situation like poker, roulette etc. These models will be based on both the domains, and will be the key to establish the relation.

Furthermore we will have 2 assignments more, plus one week in last dedicated for additional contents to cover upon discussions with mentees.

THANK YOU