Lecture 4

3 Conditional Probability and Bayesian Inference

3.1 Updation Rule (Conditional Probability)

Definitions:

- **Unconditional Probability:** $P(A)$ — Probability of event A without any condition.
- **Conditional Probability:** $P(A \mid B)$ — Probability of event A given that event B has occurred. This represents the updated belief after incorporating new information.

Example: In a manufacturing context — suppose we want to determine whether more or fewer items will be sold given the current economic condition.

3.2 Concept of Prior & Posterior

- **Prior Probability:** $P(A)$ — Initial belief about event A, before observing new evidence.
- **Posterior Probability:** $P(A \mid B)$ — Updated belief about A after observing evidence B.

Bayes' Theorem:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B \mid A)}{P(B)}$$

Bayes' theorem requires knowledge of *priors* and *likelihoods*.

Example: Manufacturing Defects from Multiple Units

Unit	Proportion of Production	Defect Probability
I	60%	0.01
II	10%	0.10
III	30%	0.20

Let:

$$D : \text{Defective item found}$$

$$I_1, I_2, I_3 : \text{Events that the item came from Unit I, II, III respectively}$$

Given:

$$P(I_1) = 0.6, \quad P(I_2) = 0.1, \quad P(I_3) = 0.3$$

$$P(D \mid I_1) = 0.01, \quad P(D \mid I_2) = 0.10, \quad P(D \mid I_3) = 0.20$$

Step 1: Compute Total Probability of Defect (Using the law of total probability)

$$P(D) = \sum_{i=1}^{3} P(I_i) \cdot P(D \mid I_i)$$

$$P(D) = (0.6)(0.01) + (0.1)(0.10) + (0.3)(0.20) = 0.006 + 0.01 + 0.06 = 0.076$$

Step 2: Compute Posterior Probabilities

Using Bayes' Theorem:

$$P(I_1 \mid D) = \frac{0.6 \cdot 0.01}{0.076} \approx 0.079$$

$$P(I_2 \mid D) = \frac{0.1 \cdot 0.10}{0.076} \approx 0.132$$

$$P(I_3 \mid D) = \frac{0.3 \cdot 0.20}{0.076} \approx 0.789$$

Although, $P(I_3) = 0.3$ (not very high) but the updated probability $P(I_3|D) = 0.789$ (maximum).

3.3 Bayes' Theorem Overview

$$P(A \mid D) = \frac{P(D \mid A) \cdot P(A)}{P(D)}$$

Where:

- $P(A \mid D)$: Posterior probability
- $P(D \mid A)$: Likelihood
- $P(A)$: Prior probability
- $P(D)$: Marginal probability of data

Applications include:

- Bayesian Networks
- Naive Bayes Classifier
- Email spam filtering

4 Independence of Events

Two events A and B are sain to be independent if:

$$P(A \cap B) = P(A) \cdot P(B)$$

If not independent:

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

Alternatively:

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

Example (Independent Events)

If:

$$P(A) = 0.8, \quad P(B) = 0.9$$

then:

$$P(A \cap B) = 0.8 \cdot 0.9 = 0.72$$

Mutual Independence of Three Events

Events A, B and C are mutually independent if:

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

$$P(A \cap B) = P(A)P(B), \quad P(B \cap C) = P(B)P(C), \quad P(A \cap C) = P(A)P(C)$$

5 Probability Theory Structure

A probability space is a triple:

$$(S, F, P)$$

where:

- S: Sample space
- F: Sigma-algebra of events
- P: Probability measure

Examples:

$$\{H, T\}, \quad \{Even, Odd\}, \quad \{1, 2, 3, 4, 5, 6\}$$

6 Random Variables

A Random Variable (RV) is:

$$X : S \to \mathbb{R}$$

Example: Coin Toss

$$S = \{H, T\}, \quad X(H) = 0, \quad X(T) = 1$$

$$P(X = 0) = \frac{1}{2}, \quad P(X = 1) = \frac{1}{2}$$

Example: Tossing a Coin 3 Times

$$S = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\}$$

Define:

$$X = \text{number of Heads}$$

Possible values:

$$X \in \{0, 1, 2, 3\},$$

$$P(X = 0) = \frac{1}{8} \quad P(X = 1) = \frac{3}{8} \quad P(X = 2) = \frac{3}{8} \quad P(X = 3) = \frac{1}{8}$$