CHAPTER – 7

Statistical Reasoning

1. Probability and Bayes' Theorem

Probability:

Probability is a mathematical framework used to quantify uncertainty. It assigns a numerical value between 0 and 1 to an event, where:

- 0 means the event will not occur.
- 1 means the event is certain to occur.
- Values in between represent varying degrees of belief.

In AI, probability is used to handle uncertain or incomplete information.

Basic Terminology:

- Sample Space (S): All possible outcomes of an experiment.
- Event (E): A subset of outcomes in the sample space.
- Probability of Event E (P(E)): Number of favorable outcomes divided by total outcomes.

Bayes' Theorem:

Bayes' Theorem provides a way to update the probability of a hypothesis based on new evidence.

Formula:

P(H|E)=P(E|H)·P(H)P(E)P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}P(H|E)=P(E)P(E|H)·P(H)

Where:

- P(H|E)P(H|E)P(H|E): Posterior probability (probability of hypothesis H given evidence E)
- P(E|H)P(E|H)P(E|H): Likelihood (probability of evidence E given H is true)
- P(H)P(H)P(H): Prior probability of hypothesis H
- P(E)P(E)P(E): Marginal probability of the evidence

Example: If a person tests positive for a disease, Bayes' theorem helps calculate the probability that the person actually has the disease, considering:

- The accuracy of the test

- The general prevalence of the disease

2. Certainty Factors and Rule-Based Systems

Certainty Factors (CF):

Certainty factors were introduced in the MYCIN expert system for medical diagnosis to manage uncertainty in rule-based systems.

A certainty factor is a number between -1 and 1 used to represent belief:

- +1 = absolute certainty that a fact is true.
- -1 = absolute certainty that a fact is false.
- 0 = no information or neutral belief.

Combining Certainty Factors:

When multiple pieces of evidence support or refute a conclusion, their CFs are combined using specific rules.

- **If two rules support the same conclusion:**

CFcombined=CF1+CF2·(1−CF1)CF_{combined} = CF_1 + CF_2 \cdot (1 - CF_1)CFcombined =CF1+CF2·(1−CF1)

- **If they refute:**

CFcombined=CF1+CF2·(1+CF1)CF_{combined} = CF_1 + CF_2 \cdot (1 + CF_1)CFcombined =CF1+CF2·(1+CF1)

Rule-Based Systems:

These systems use IF-THEN rules to derive conclusions:

- IF condition THEN conclusion (CF = value) They are simple and intuitive but can struggle with conflicting or uncertain information. Certainty Factors help mitigate this issue.

3. Bayesian Networks

Definition:

A Bayesian Network (BN) is a graphical model representing a set of variables and their conditional dependencies via a directed acyclic graph (DAG).

Each node in the graph represents a variable, and each edge denotes a dependency between variables.

Components:

- Nodes = Variables (random variables, e.g., "Fever", "Flu")
- Edges = Dependencies (e.g., Flu → Fever)
- Conditional Probability Tables (CPTs): Define the probability of each node given its parent(s).

Advantages:

- Efficient handling of uncertainty
- Clear visualization of dependencies
- Supports both inference (calculating probabilities) and learning (from data)

Example:

Let’s say we have a Bayesian network:

Flu → Fever

Cold → Cough

Flu, Cold → Fatigue

If we observe Fever and Fatigue, we can infer the probability that the person has the Flu or Cold using Bayesian inference.

4. Dempster-Shafer Theory (Evidence Theory)

Overview:

Dempster-Shafer Theory is a generalization of Bayesian probability theory. It allows for representation of ignorance and combination of evidence from different sources.

Key Concepts:

- Frame of Discernment (Θ): The set of all possible hypotheses.
- Basic Probability Assignment (BPA): Assigns probability mass to sets of hypotheses, not just individual ones.
- Belief Function (Bel): Degree of belief supporting a hypothesis.
- Plausibility Function (Pl): Degree to which a hypothesis is not refuted by evidence.

$$Bel(A) \leq Pl(A)$$

Dempster's Rule of Combination:

Used to combine two BPAs (from different sources) into a new one.

$$m(C) = \frac{1}{1 - K} \sum_{A \cap B = C} m_1(A) \cdot m_2(B)$$

Where:

- m_1, m_2 are BPAs from different sources.
- K is the conflict between the sources.

Why Use It?

Unlike Bayesian theory, Dempster-Shafer can:

- Represent total ignorance (e.g., assign mass to the whole set Θ)
- Combine evidence without requiring precise probabilities

Applications:

- Sensor fusion
- Fault diagnosis
- Decision making under uncertainty

Summary Table

Topic	Key Idea	Method Used	Application
Probability	Quantify uncertainty	Classical/Conditional	Prediction, diagnostics
Bayes' Theorem	Update beliefs	Prior + Evidence	Disease diagnosis, spam filters
Certainty Factors	Degree of belief	CF calculation	Expert systems
Bayesian Networks	Graphical probabilistic model	DAG + CPTs	Decision making, AI
Dempster-Shafer	Combine uncertain evidence	BPA + Bel/Pl functions	Multisource reasoning