Statistical Decision Theory

ANISHA JOSEPH

Statistical Decision Theory

Statistical Decision Theory is a framework used to make decisions in the presence of uncertainty. It combines statistical inference with decision-making under uncertainty.

In hypothesis testing, we often need to make decisions about which hypothesis to accept, based on observed data.

The two key concepts here—***Complete Class of Decision Rules*** **and** ***Minimal Complete Class of Decision Rules***—are related to the optimality of decision rules in a statistical decision-making context.

Statistical Decision Theory

Decision Rule: A rule that determines which hypothesis (either $H0$ or $H1$) to accept based on the observed data.

Loss Function: Represents the loss (or cost) incurred when a particular decision is made. For example, rejecting $H0$ when it is true (Type I error) or failing to reject $H0$ when $H1$ is true (Type II error).

Risk Function: The expected loss, given a decision rule, which takes into account both the loss function and the probability distribution of the data.

Complete Class of Decision Rules

A **complete class** is a set of decision rules that are "good" in the sense that no rule outside this set can perform better in terms of the expected loss (risk function).

In other words, a decision rule belongs to a complete class if it is at least as good as any other rule that doesn't belong to the class.

In the context of hypothesis testing, a complete class includes all the decision rules that are optimal in the sense that no other rules can result in a lower expected loss.

Minimal Complete Class of Decision Rules:

A **minimal complete class** is the **smallest subset of the complete class** that contains all the decision rules that are optimal with respect to the loss function.

It is a "minimal" set in the sense that if we remove any rule from it, we would no longer have a complete class.

In hypothesis testing, the minimal complete class is essentially the subset of decision rules that cannot be improved upon, even within the complete class.

Example

Example for Z test:

Likelihood Ratio Test (Z-test) Rule: based on calculated Z value

Alternative decision rules:

P-value-based Decision Rule:

Confidence Interval-based Decision Rule:

Optimal Decision Rule

Definition: An optimal decision rule minimizes the expected loss or risk in making decisions under uncertainty.

Key Concept: The goal is to balance Type I error (false positive) and Type II error (false negative) in a way that minimizes overall risk.

Risk Function: Combines the probability of error and the consequences of those errors.

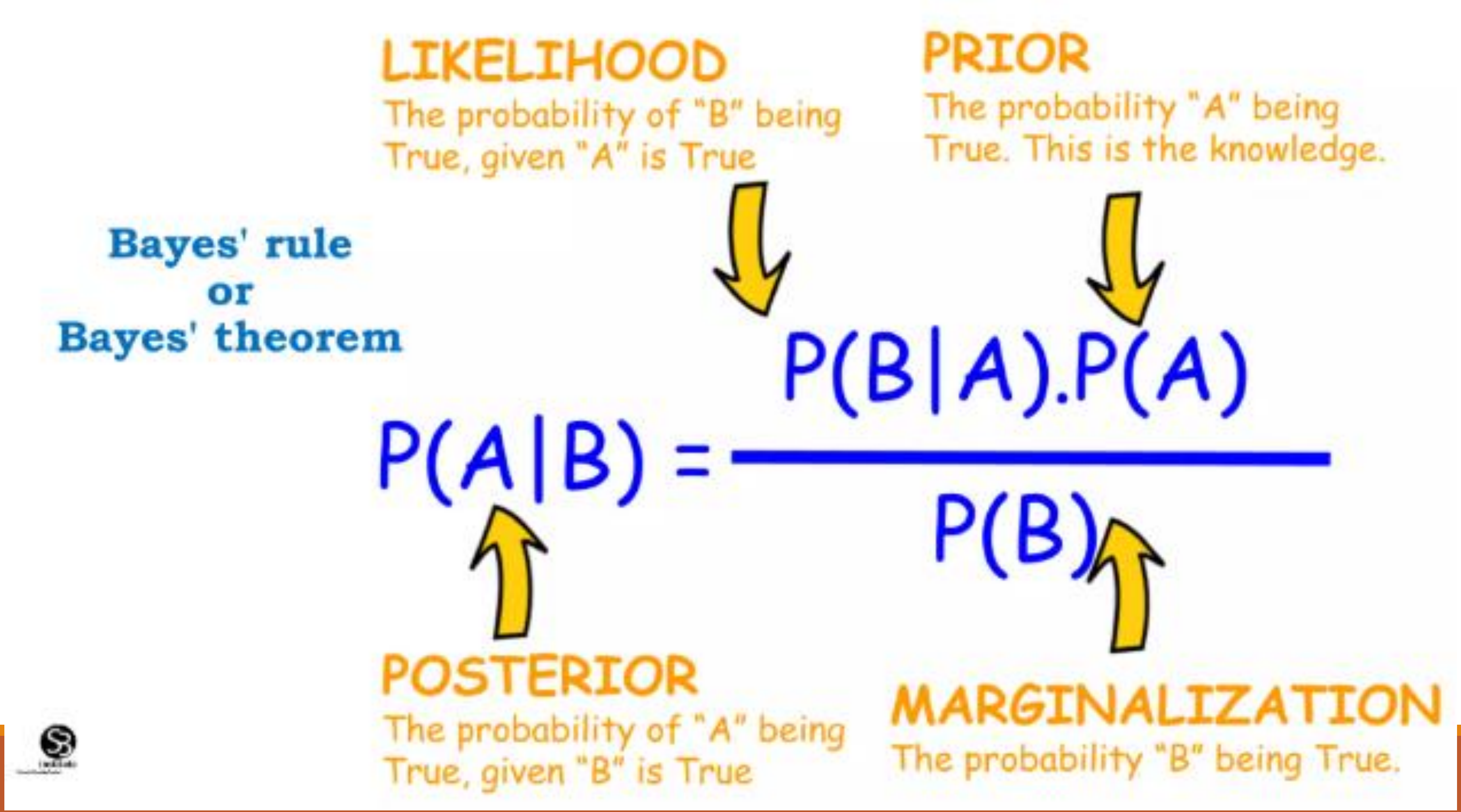
Method of Finding a Bayes Rule

Bayes' theorem

- Bayes' theorem can be derived using product rule and conditional probability of event A with known event B:
- As from product rule we can write:
- $P(A \wedge B) = P(A|B)\,P(B)$ or
- imilarly, the probability of event B with known event A:
- $P(A \wedge B) = P(B|A)\,P(A)$
- Equating right hand side of both the equations, we will get:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \qquad \text{....(a)}$$

- The above equation (a) is called as **Bayes' rule** or **Bayes' theorem**. This equation is basis of most modern AI systems for **probabilistic inference**.


LIKELIHOOD
The probability of "B" being True, given "A" is True
PRIOR
The probability "A" being True. This is the knowledge.
Bayes' rule
or
Bayes' theorem
$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$
POSTERIOR
The probability of "A" being True, given "B" is True
MARGINALIZATION
The probability "B" being True.

A Statistician travels by local train to his music coaching classes which are held for a maximum of 5 days a week. However the classes are held at random and he is informed one day prior if classes are scheduled for the next day.

His daily trip cost one way to his destination is \$ 50. He needs to decide if he wishes to purchase a weekly train pass which will cost him \$ 320. The probability distribution of the number of times coaching classes are held in a week are:

Number of classes held in a week	Probability
1	0.2
2	0.1
3	0.3
4	0.25
5	0.15

Given that he never misses a scheduled class and travels by train only, find out his minimax and Bayes criterion decisions.

Solution

Part (a)

•

The costs incurred by the student attending coaching classes depending on the number of classes are:

No of classes	Cost(to and fro) from coaching class (in $)
1	100
2	200
3	300
4	400
5	500

If he purchases the weekly train pass, the cost incurred equals $ 320

- d1 = decision of not buying the weekly pass
- d2 = decision of buying the weekly pass

Minimax criterion:

- The maximum cost under the decision (d1) of not purchasing weekly train pass is $ 500
- The maximum cost under the decision (d2) of purchasing weekly train pass is $ 320
- The minimax solution is the minimum of the maximum costs under the two decision functions i.e. decision d2

- Thus under the minimax criterion the student should choose to buy a weekly train pass

Bayes Solution:

- The decision which would minimise the expected cost would be the Bayes solution to this problem.
- Under d1, the expected cost equals

$$\begin{aligned} &= 100 \times 0.2 + 200 \times 0.1 + 300 \times 0.3 + 400 \times 0.25 + 500 \times 0.15 \\ &= 20 + 20 + 90 + 100 + 75 \\ &= 305 \end{aligned}$$

- Under d2, the expected cost is the cost of purchasing the pass, which is equal to 320

- Thus d1 (i.e. not purchasing a weekly train pass), is the decision which minimises the expected cost and hence is the Bayes Solution to this problem.

A manufacturer of specialist products for the retail market must decide which product to make in the coming year.

- There are three possible choices basic, Deluxe or Supreme, each with different tooling up costs.
- The manufacturer has fixed overheads of $1,300,000.

The revenue and tooling up costs for each product are as follows:

	Tooling up costs	Revenue per item sold
Basic	100,000	1.00
Deluxe	400,000	1.20
Supreme	1,000,000	1.50

Last year the manufacturer sold 2,100,000 items and is preparing forecasts of profitability for the coming year based on three scenarios:

- Low sales (70% of last year's level);
- Medium sales (same as last year)
- High sales (15% higher than last year).

(a) Determine the annual profits in $ under each possible combination.

(b) Determine the minimax solution to this problem.

(c) Determine the Bayes criterion solution based on the annual profit given the probability distribution: $P(Low) = 0.25$; $P(Medium) = 0.6$ and $P(High) = 0.15$.

Solutions

Part (a)

Revenue ($)

Last year's revenue 2,100,000 scaled by a combination of price and market condition

	Low (0.70)	Medium (1.00)	High (1.15)
Basic (1.00)	1.00×0.70	1.00×1.00	1.00×1.15
Deluxe (1.20)	1.20×0.70	1.20×1.00	1.20×1.15
Supreme (1.50)	1.50×0.70	1.50×1.00	1.50×1.15

	Low	Medium	High
Basic	1,470,000	2,100,000	2,415,000
Deluxe	1,764,000	2,520,000	2,898,000
Supreme	2,205,000	3,150,000	3,622,500