Bayesian Decision Theory is a fundamental statistical approach to pattern recognition and classification. It leverages Bayes' theorem to make decisions based on the probabilities of different outcomes and their associated costs. Here’s a **step-by-step guide to using Bayesian Decision Theory for object classification**:

Step 1: Define the Classes

Assume we have NNN classes C1,C2,...,CNC_1, C_2, ..., C_NC1,C2,...,CN. Each class represents a different object category we want to classify.

Step 2: Determine the Prior Probabilities
The prior probability P(Ci)P(C_i)P(Ci) represents the probability that an object belongs to class CiC_iCi before any observations are made.

Step 3: Likelihood Function
The likelihood function P(x∣Ci)P(x | C_i)P(x∣Ci) is the probability of observing data xxx given that the object belongs to class CiC_iCi.

Step 4: Bayes' Theorem

Bayes' theorem combines the prior probability and the likelihood function to calculate the posterior probability $P(C_i|x)$, which is the probability that the object belongs to class Ci given the observed data

$$P(C_i|x) = \frac{P(x|C_i)P(C_i)}{P(x)}$$

where $P(x) = \sum_{j=1}^{N} P(x|C_j)P(C_j)$.

Step 5: Decision Rule

The decision rule assigns the object to the class with the highest posterior probability. For each class C_i, calculate:

$$P(C_i|x)$$

Then, classify x to the class C_k such that:

$$k = \arg\max_i P(C_i|x)$$

Example : **To classify an object as either belonging to Class 1 or Class 2 based on a feature xxx using Bayesian Decision Theory in simple terms:**

Step 1: Define the Classes

We have two classes:

- Class 1 (C1)
- Class 2 (C2)

Step 2: Determine the Prior Probabilities

The prior probabilities tell us the likelihood of each class before we look at the feature x:

- $P(C1)$: Probability that the object belongs to Class 1.
- $P(C2)$: Probability that the object belongs to Class 2.

Let's assume:

- $P(C1) = 0.6$
- $P(C2) = 0.4$

Step 3: Likelihood Function

The likelihood function gives the probability of observing a feature x given that the object is from a specific class:

- $P(x|C1)$: Probability of observing x if the object is from Class 1.
- $P(x|C2)$: Probability of observing x if the object is from Class 2.

Step 4: Bayes' Theorem

We use Bayes' theorem to update our belief about the class of the object based on the observed feature x:

$$P(C1|x) = \frac{P(x|C1)P(C1)}{P(x)}$$

$$P(C2|x) = \frac{P(x|C2)P(C2)}{P(x)}$$

Where $P(x)$ is the total probability of observing x:

$$P(x) = P(x|C1)P(C1) + P(x|C2)P(C2)$$

Step 5: Decision Rule

We classify the object based on which class has the higher posterior probability:

- If $P(C1|x) > P(C2|x)$, classify the object as Class 1.
- If $P(C2|x) > P(C1|x)$, classify the object as Class 2.

Example

Let's classify an object with feature x=6x = 6x=6.

1.Prior Probabilities:

1. P(C1)=0.6P(C1) = 0.6P(C1)=0.6
2. P(C2)=0.4P(C2) = 0.4P(C2)=0.4

2.Likelihood Functions (assuming Gaussian distributions):

- $P(x|C1) = \mathcal{N}(x; \mu_1 = 5, \sigma_1 = 1)$
- $P(x|C2) = \mathcal{N}(x; \mu_2 = 7, \sigma_2 = 1.5)$

Likelihood Functions

For Class 1 (C1C_1C1):

- The likelihood function P(x∣C1)P(x | C_1)P(x∣C1) describes the probability of observing the feature xxx given that the object belongs to Class 1.
- We assume this probability follows a Gaussian (normal) distribution with:
    - Mean (μ1\mu_1μ1) = 5
    - Standard deviation (σ1\sigma_1σ1) = 1

In simple terms:

- If an object is from Class 1, its feature xxx is most likely to be around 5, but can vary within a range, typically not too far from 5.

For Class 2 (C2C_2C2):

- The likelihood function P(x∣C2)P(x | C_2)P(x∣C2) describes the probability of observing the feature xxx given that the object belongs to Class 2.
- We assume this probability also follows a Gaussian (normal) distribution but with:
    - Mean (μ2\mu_2μ2) = 7
    - Standard deviation (σ2\sigma_2σ2) = 1.5

In simple terms:

- If an object is from Class 2, its feature xxx is most likely to be around 7, but can vary within a range, typically not too far from 7.

Visual Representation

To visualize:

- Class 1's feature xxx clusters around 5, with most values falling within 4 to 6.
- Class 2's feature xxx clusters around 7, with most values falling within 5.5 to 8.5.

Using these likelihood functions, we can calculate the probabilities of observing a specific feature value xxx for each class, helping us classify the object correctly.

3. **Calculate Likelihoods:**

$$P(6|C1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(6-5)^2}{2}}$$

$$P(6|C2) = \frac{1}{\sqrt{2\pi \cdot 1.5^2}} e^{-\frac{(6-7)^2}{2 \cdot 1.5^2}}$$

4. **Calculate Posterior Probabilities:**

$$P(C1|6) \propto P(6|C1) \times P(C1)$$

$$P(C2|6) \propto P(6|C2) \times P(C2)$$

Since $P(x)$ is a normalizing constant, we can compare the numerators directly:

$$P(C1|6) \propto \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}} \times 0.6$$

$$P(C2|6) \propto \frac{1}{\sqrt{2\pi \cdot 2.25}} e^{-\frac{1}{4.5}} \times 0.4$$

5. **Compare Posterior Probabilities**:

- Calculate the values and compare:
  - If $P(C1|6) > P(C2|6)$, classify as Class 1.
  - If $P(C2|6) > P(C1|6)$, classify as Class 2.

For simplicity, let's assume the calculations show $P(C1|6) > P(C2|6)$. Therefore, we classify the object as belonging to Class 1.