Guide to Probability Concepts in AI & ML

# 1. Classical Probability

Definition:  
Classical probability is the measure of the likelihood of an event happening, assuming all outcomes are equally likely.

Formula:  
P(E) = Number of favorable outcomes / Total number of possible outcomes

Example:  
What is the probability of getting a head when tossing a fair coin?

Favorable outcomes = 1 (Head)  
Total outcomes = 2 (Head, Tail)

P(Head) = 1 / 2

# 2. Conditional Probability

Definition:  
The probability of event A occurring given that event B has already occurred.

Formula:  
P(A|B) = P(A ∩ B) / P(B), provided P(B) > 0

Example:  
P(A ∩ B) = 0.3, P(B) = 0.6 => P(A|B) = 0.3 / 0.6 = 0.5

# 3. Independent Events

Definition:  
Two events A and B are independent if the occurrence of one does not affect the probability of the other.

Formula:  
P(A ∩ B) = P(A) × P(B)

Also, P(A|B) = P(A), P(B|A) = P(B)

Example:  
Toss a coin and roll a die: P(Head) = 0.5, P(rolling a 6) = 1/6 => P(Head and 6) = 0.5 × 1/6 = 1/12

# 4. Total Probability Theorem

Definition:  
Gives a way to compute the probability of an event based on several different mutually exclusive conditions.

Formula:  
P(A) = ∑ P(A|Bi) × P(Bi), where {B1, B2, ..., Bn} partitions the sample space

Example:  
Two populations:

- Population 1: 60% → P(Pos|P1) = 0.05

- Population 2: 40% → P(Pos|P2) = 0.10

P(Positive) = (0.05 × 0.6) + (0.10 × 0.4) = 0.03 + 0.04 = 0.07

# 5. Bayes’ Theorem

Definition:  
Allows updating the probability of an event based on new evidence.

Formula:  
P(Bi|A) = [P(A|Bi) × P(Bi)] / ∑ [P(A|Bj) × P(Bj)]

For two events A and B: P(B|A) = [P(A|B) × P(B)] / P(A)

Example (Disease Testing):

- P(Disease) = 0.01, P(Pos|D) = 0.95, P(Pos|¬D) = 0.05

- P(¬D) = 0.99

- P(Pos) = (0.95×0.01) + (0.05×0.99) = 0.0095 + 0.0495 = 0.059

- P(D|Pos) = (0.95×0.01)/0.059 ≈ 0.161 → 16.1% chance of actually having disease

# 6. Applications in AI & ML

Classical Probability: Used in randomized algorithms and probabilistic models.

Conditional Probability: Naive Bayes, Markov Chains, NLP tasks.

Independent Events: Assumed in Naive Bayes classifier.

Total Probability: Used in ensemble learning and model selection.

Bayes’ Theorem: Core to Bayesian Networks, Naive Bayes Classifier, Predictive Modeling.

# 7. Summary of Formulas

Classical Probability: P(E) = Favorable / Total

Conditional Probability: P(A|B) = P(A ∩ B) / P(B)

Independent Events: P(A ∩ B) = P(A) × P(B)

Total Probability: P(A) = ∑ P(A|Bi) × P(Bi)

Bayes’ Theorem: P(B|A) = [P(A|B) × P(B)] / P(A)