Probability & Probability Distributions – ML/DS Tasks

# Task 1: Coin Toss Simulation for Classification

Simulate tossing a coin 1000 times and estimate the probability of heads. Then relate how this is similar to predicting probabilities in a binary classification problem (e.g., spam vs. not spam).

# Task 2: Dice Roll & Feature Engineering

Roll a dice 10,000 times and calculate the probability of each outcome. Discuss how this relates to categorical feature encoding in machine learning.

# Task 3: Conditional Probability in Data Filtering

You are analyzing a dataset of customers.  
- 40% of customers purchased a product.  
- 20% of all customers are students.  
- 10% are students and purchased the product.  
Calculate P(Purchase | Student). Relate this to targeted marketing in ML models.

# Task 4: Normal Distribution in Model Errors

Generate random values from a normal distribution (mean=0, std=1) to simulate model errors. Explain why many ML models assume errors follow a Gaussian distribution.

# Task 5: Uniform Distribution in Neural Networks

Simulate sampling from a uniform distribution to initialize model weights. Explain why uniform initialization is used in deep learning.

# Task 6: Bayesian Probability in Medical Diagnosis

In a dataset:  
- 1% of patients have a rare disease.  
- A test detects the disease correctly 99% of the time.  
- False positive rate = 5%.  
Using Bayes’ theorem, calculate the probability that a patient actually has the disease given a positive test. Explain its importance in medical AI models.

# Task 7: Probability Distribution Visualization

Take a real dataset (e.g., Titanic dataset from Kaggle). Plot the distribution of passenger ages.  
- Check if it is closer to a normal distribution or skewed distribution.  
- Explain why checking data distribution is important before applying ML algorithms.

# Task 8: Probability Mass Function (PMF) in NLP

Take a text dataset (any article). Count the frequency of each word and normalize it to get probabilities. This is the PMF of words. Explain how this concept is used in Naive Bayes for text classification.