

## **Statistics Advanced – 1**

Question 1: What is a random variable in probability theory?

- A random variable in probability theory is a function that assigns a numerical value to each outcome in a sample space of a random experiment. Its value is determined by chance, making it inherently uncertain until the experiment is conducted.

Question 2: What are the types of random variables?

- Random variables in probability theory are classified mainly into two types based on the nature of their possible values: discrete and continuous. Less commonly, mixed types combine both characteristics.

Question 3: Explain the difference between discrete and continuous distributions.

- **1. Discrete Distribution:**

Definition: A discrete distribution describes the probability of outcomes for a discrete random variable. A discrete random variable can take on only a finite or countably infinite set of distinct values (like 0, 1, 2, 3, etc.).

Examples: The number of heads in a series of coin flips, the number of cars passing through a toll booth, or the number of calls received at a call center.

Probability Mass Function (PMF): The probabilities are assigned to individual outcomes. The sum of all probabilities in a discrete distribution is always 1.

Example Distribution: The binomial distribution, Poisson distribution, and geometric distribution are examples of discrete distributions.

Key Characteristic: For a discrete distribution, the random variable takes specific, isolated values.

- **2. Continuous Distribution:**

Definition: A continuous distribution describes the probability of outcomes for a continuous random variable, which can take any value within a given range or interval (for example, any real number between 0 and 1, or any value within a specific time range).

Examples: The height of a person, the time it takes to run a marathon, the temperature on a given day.

Probability Density Function (PDF): Rather than assigning probabilities to specific outcomes, a continuous distribution assigns probabilities to ranges of values. The area under the curve of the probability density function (PDF) over a given interval represents the probability of the variable falling within that range.

**Example Distribution:** The normal distribution, exponential distribution, and uniform distribution (when defined on a continuous interval) are examples of continuous distributions.

**Question 4:** What is a binomial distribution, and how is it used in probability?

- A binomial distribution is a type of discrete probability distribution that describes the number of successes in a fixed number of independent trials, where each trial has only two possible outcomes.

**Question 5:** What is the standard normal distribution, and why is it important?

- Standard Normal Distribution
- The standard normal distribution is a special case of the normal distribution that has:
  - Mean ( $\mu$ ) = 0
  - Standard deviation ( $\sigma$ ) = 1
  - It is commonly denoted as:
  - It is a continuous probability distribution that is symmetric and bell-shaped.

**Question 6:** What is the Central Limit Theorem (CLT), and why is it critical in statistics?

- The Central Limit Theorem (CLT) states that the sampling distribution of the sample mean approaches a normal distribution as the sample size increases (typically  $n \geq 30$ ), regardless of the population's underlying distribution, provided the samples are independent and identically distributed with finite variance. This holds under mild conditions, making the mean approximately normally distributed with mean  $\mu$  (population mean) and standard deviation  $\sigma/\sqrt{n}$ .

**Question 7:** What is the significance of confidence intervals in statistical analysis?

- Confidence intervals (CIs) provide a range of plausible values for an unknown population parameter, such as a mean or proportion, based on sample data, quantifying the uncertainty around a point estimate like a sample mean.

**Question 8:** What is the concept of expected value in a probability distribution?

- The expected value of a probability distribution, often denoted  $[E(X)]$  or  $[\mu]$ , is the long-term average outcome of a random variable  $[X]$  if the experiment is repeated infinitely many times. It serves as the mean or center of mass of the distribution, weighted by the probabilities of each possible value.