Statistics Advance Part 1

1. What is a random variable in probability theory?

Ans- A random variable is a numerical function that assigns a real number to each outcome of a random experiment. Its value depends on chance.

Example: Let X be the number of heads obtained when tossing a coin three times.

2. What are the types of random variables?

Asn- There are two types of random variables:

1. **Discrete random variable** – Takes countable values (e.g., number of students).
2. **Continuous random variable** – Takes values from a continuous range (e.g., height, weight).

3. What is the difference between discrete and continuous distributions?

Ans-

- Discrete distributions deal with countable values and use a probability mass function (PMF).
- Continuous distributions deal with uncountable values and use a probability density function (PDF).
- In discrete distributions, probabilities are assigned to exact values; in continuous distributions, probabilities are assigned over intervals.

4. What are probability distribution functions (PDF)?

Ans- A probability distribution function (PDF) describes how probability is distributed over values of a random variable.

- For discrete variables, it gives the probability of each value
- For continuous variables, it gives the probability density, and total area under the curve equals 1.

5. How do cumulative distribution functions (CDF) differ from probability distribution functions (PDF)?

Ans-

- PDF gives the probability (or density) at a specific value.
- CDF gives the probability that the random variable is less than or equal to a given value.
- CDF is obtained by summing or integrating the PDF.

6. What is a discrete uniform distribution?

Ans- A discrete uniform distribution is one where all possible discrete outcomes have equal probability.

Example: Outcomes of a fair die, where each number has probability ⅙

7. What are the key properties of a Bernoulli distribution?

Ans- Only two outcomes: success (1) and failure (0).
Probability of success = ppp, failure = 1−p1-p1−p.
Mean = ppp, Variance = p(1−p)p(1-p)p(1−p).
It is the simplest discrete probability distribution.

8. What is the binomial distribution, and how is it used in probability?

Ans- The binomial distribution gives the probability of getting a certain number of successes in n independent trials, each with success probability ppp. It is used in experiments like coin tosses, quality testing, and surveys.

9. What is the Poisson distribution and where is it applied?

Ans- The Poisson distribution models the number of events occurring in a fixed interval of time or space when events happen independently at a constant rate.

Applications include call arrivals, accidents, and system failures.

10. What is a continuous uniform distribution?

Ans- A continuous uniform distribution assigns equal probability density to all values within a given interval [a,b]

Its mean is a+b/2 and it is used when all outcomes are equally likely.

11. What are the characteristics of a normal distribution?

Ans-

- Bell-shaped and symmetric
- Mean = Median = Mode
- Defined by mean μ\muμ and standard deviation σ\sigmaσ
- Total area under the curve equals 1

12. What is the standard normal distribution, and why is it important?

Ans- A standard normal distribution has mean 0 and standard deviation 1.
It is important because it allows comparison of different datasets using Z-scores.

13. What is the Central Limit Theorem (CLT), and why is it critical in statistics?

Ans- The CLT states that the distribution of sample means approaches a normal distribution as sample size increases, regardless of the population distribution.It is critical because it enables statistical inference.

14. How does the Central Limit Theorem relate to the normal distribution?

Ans- CLT explains why the normal distribution appears frequently in real data. Even if the original data is not normal, the sample mean will follow a normal distribution for large samples.

15. What is the application of Z statistics in hypothesis testing?

Ans- Z statistics are used to test hypotheses about population means when population variance is known and sample size is large.They help decide whether to reject or accept the null hypothesis.

16. How do you calculate a Z-score, and what does it represent?

Ans- $Z= x-\mu/ \sigma$

A Z-score represents how many standard deviations a value is away from the mean.

17.What are point estimates and interval estimates in statistics?

Ans- Point estimate: A single value used to estimate a population parameter (e.g., sample mean).

Interval estimate: A range of values within which the parameter is expected to lie (e.g., confidence interval)

18. What is the significance of confidence intervals in statistical analysis?

Ans- Confidence intervals provide a range of plausible values for a population parameter and indicate the reliability of an estimate.

19. What is the relationship between a Z-score and a confidence interval?

Ans- Confidence intervals are constructed using Z-scores, which determine the margin of error based on the confidence level.

20. How are Z-scores used to compare different distributions?

Ans- Z-scores standardize data, allowing comparison of values from different distributions with different means and standard deviations.

21. What are the assumptions for applying the Central Limit Theorem?

Ans-

- Sample size is sufficiently large
- Observations are independent
- Population has finite mean and variance

22. What is the concept of expected value in a probability distribution?

Ans- The expected value is the long-run average value of a random variable. It represents the center or mean of the distribution.

23. How does a probability distribution relate to the expected outcome of a random variable?

Ans- The probability distribution assigns probabilities to outcomes, and the expected value is calculated as a weighted average of these outcomes using their probabilities.