Statistics Advance-1

1. What is a random variable in probability theory?

A **random variable** is a variable whose possible values are numerical outcomes of a random phenomenon. It maps outcomes of a random experiment to numbers.

2. What are the types of random variables?

- **Discrete Random Variables**: Take countable values (e.g., 0, 1, 2).
- **Continuous Random Variables**: Take any value within a given range or interval (e.g., 2.31, 5.67).

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

- **Discrete distribution** describes the probability of outcomes of a discrete random variable.
- **Continuous distribution** describes probabilities for continuous variables using areas under a curve.

4. What are probability distribution functions (PDF)?

PDF (Probability Density Function) describes the **likelihood of a continuous random variable** taking on a specific value. The area under the PDF curve over an interval gives the probability of that interval.

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

CDF gives the **probability that a variable is less than or equal to a value**.

- PDF shows probability **density** at a point.
- CDF is the **cumulative sum** of probabilities up to a point.

6. What is a discrete uniform distribution?

A **discrete uniform distribution** is one where each possible value has an equal probability. Example: rolling a fair die (each outcome from 1 to 6 has probability 1/6).

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

- Two outcomes: success (1) or failure (0)
- One trial

Statistics Advance-1

- Probability of success = p, failure = 1 − p
- Mean = p, Variance = p(1 − p)

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

Describes the number of successes in **n independent Bernoulli trials** with the same probability of success (p).
Used in scenarios like coin tosses, quality checks, etc.

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

Models the **number of events occurring in a fixed interval** of time or space, given that they occur with a known average rate (λ).
Used in: call arrivals, traffic flow, server requests.

10. What is a continuous uniform distribution?

A distribution where all outcomes in a continuous interval are **equally likely**.
PDF is constant within the interval [a, b]:
f(x)=1b−afor a≤x≤bf(x) = \frac{1}{b - a} \quad \text{for } a \le x \le bf(x)=b−a1for a≤x≤b

11. What are the characteristics of a normal distribution?

- Bell-shaped, symmetric around the mean
- Mean = Median = Mode
- Defined by mean (μ) and standard deviation (σ)
- 68%-95%-99.7% rule applies

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

A normal distribution with **mean = 0** and **standard deviation = 1**.
Used to compute probabilities using **Z-scores** and standard normal tables.

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

CLT states that the sampling distribution of the sample mean **approaches a normal distribution** as sample size increases, regardless of the population's distribution.
Important for making inferences and hypothesis testing.

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

Statistics Advance-1

CLT explains why the **normal distribution is so widely used**—it justifies normal approximation of sample means and proportions for large sample sizes.

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

Z-statistics are used when population variance is known and sample size is large to determine if the sample mean significantly differs from the population mean.

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

Z=X−μσZ = \frac{X - \mu}{\sigma}Z=σX−μ

Represents how many standard deviations a value (X) is from the mean (μ).

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

- **Point estimate**: Single value estimate of a population parameter (e.g., sample mean).
- **Interval estimate**: Range of values (confidence interval) where the parameter likely lies.

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

They provide a **range of plausible values** for a population parameter and reflect **uncertainty** in estimation.

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

Z-scores determine the **margin of error** in a confidence interval (e.g., 1.96 for 95% confidence).
CI = Point Estimate ± (Z * Standard Error)

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

Z-scores standardize different datasets, allowing comparison by converting values to a **common scale** based on mean and standard deviation.

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

- Random, independent sampling
- Large enough sample size (typically $n \geq 30$)
- Finite population variance

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

Statistics Advance-1

The **long-run average** value of a random variable.

E(X)=∑[xi⋅P(xi)](for discrete)E(X) = \sum [x_i \cdot P(x_i)] \quad \text{(for discrete)}E(X)=∑[xi⋅P(xi)](for discrete) E(X)=∫x⋅f(x)dx(for continuous)E(X) = \int x \cdot f(x) dx \quad \text{(for continuous)}E(X)=∫x⋅f(x)dx(for continuous)

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

The **shape and values of the distribution** determine the **likelihood** of outcomes and the **expected value** gives a central measure of those outcomes.