SAMPLING DISTRIBUTION
CORE – V

SHORT TYPE QUESTIONS.

1. Let X1, X2, ..., Xn be a sequence of independent random variables with mean $\mu = 10$ and variance $\sigma^2 = 25$. If n = 100, what is the mean of the sample mean $\bar{X}$?
2. A random variable X follows a Poisson distribution with parameter $\lambda = 3$. Find the probability $P(X \leq 5)$.
3. Let Y1, Y2, ..., Yn be a sequence of random variables with mean $\mu = 50$ and variance $\sigma^2 = 16$. If n = 25, what is the variance of the sample mean $\bar{Y}$?
4. The weight of a product follows a normal distribution with mean 500 grams and standard deviation 10 grams. What is the probability that the weight is between 490 and 510 grams?
5. Consider a sequence of Bernoulli random variables X1, X2, ..., Xn with success probability p = 0.6. If n = 50, what is the variance of the sample mean $\bar{X}$?
6. A coin is flipped 100 times, and the number of heads follows a binomial distribution. If the probability of getting a head is 0.4, find the probability of getting exactly 35 heads.
7. Let Z1, Z2, ..., Zn be a sequence of independent standard normal random variables. Find the probability $P(Z > 1.96)$.
8. The annual rainfall in a city follows an exponential distribution with a mean of 60 inches. What is the probability that the rainfall in a particular year will be less than 70 inches?
9. A random variable X follows a uniform distribution between 0 and 1. Find the mean of X.
10. Let X1, X2, ..., Xn be a sequence of independent random variables with mean $\mu = 25$ and variance $\sigma^2 = 36$. If n = 64, what is the standard deviation of the sample mean $\bar{X}$?
11. The ________________ states that for a sufficiently large n, the distribution of the sample mean $\bar{X}$ of n i.i.d. random variables approaches a normal distribution with mean μ and variance σ^2/n.
12. Chebyshev's inequality states that for any random variable X with mean μ and variance σ^2, the probability $P(|X-\mu| \geq k\sigma)$ is ________________ or less, where k is any positive constant.
13. The Weak Law of Large Numbers (W.L.L.N.) states that the sample mean $\bar{X}$ of a large number of i.i.d. random variables approaches the ________________ as the sample size n approaches infinity.
14. The De-Moivre Laplace theorem states that as the number of trials n approaches infinity, the binomial distribution with probability of success p converges to the ________________ distribution.
15. The Central Limit Theorem (C.L.T.) applies to the sum or average of a large number of ________________ random variables, irrespective of the shape of the original distribution.
16. The probability density function of the standard normal distribution is given by f(z) = ________________, where φ represents the standard normal distribution.
17. The convergence in probability of a sequence of random variables Xn to a random variable X means that for any $\varepsilon > 0$, the probability $P(|Xn - X| \geq \varepsilon)$ approaches ________________ as n approaches infinity.
18. The sample mean $\bar{X}$ is an unbiased estimator of the population ________________ μ.
19. The variance of the sample mean $\bar{X}$ is equal to the population variance σ^2 divided by the sample size n, denoted as ________________.
20. The probability distribution of the number of successes in a fixed number of independent Bernoulli trials with constant probability of success p is given by the ________________ distribution.
21. The Central Limit Theorem (C.L.T.) applies to the sum or average of a large number of
    a) Independent random variables
    b) Dependent random variables
    c) Non-identically distributed random variables
    d) Uniformly distributed random variables
22. The Weak Law of Large Numbers (W.L.L.N.) states that the sample mean $\bar{X}$ of a large number of i.i.d. random variables approaches
    a) Zero
    b) The sample mean of the first few observations
    c) The sample mean of the last few observations
    d) The population mean

23. Chebyshev's inequality states that for any random variable X with mean μ and variance σ^2, the probability $P(|X-\mu| \geq k\sigma)$ is
    a) Less than k
    b) Greater than k
    c) Equal to k
    d) Less than or equal to $1/k^2$
24. The probability density function of the standard normal distribution is given by
    a) $f(z) = e^{-z^2}$
    b) $f(z) = 1 - e^{-z}$
    c) $f(z) = (1/\sqrt{2\pi}) * e^{-z^2/2}$
    d) $f(z) = z * e^{-z}$
25. The De-Moivre Laplace theorem states that as the number of trials n approaches infinity, the binomial distribution with probability of success p converges to the
    a) Exponential distribution
    b) Poisson distribution
    c) Gamma distribution
    d) Normal distribution
26. Convergence in probability of a sequence of random variables Xn to a random variable X means that for any $\varepsilon > 0$, the probability $P(|X_n - X| \geq \varepsilon)$ approaches
    a) 0
    b) 1
    c) ε
    d) Infinity
27. The variance of the sample mean $\bar{X}$ is equal to the population variance σ^2 divided by the sample size n, denoted as
    a) n/σ^2
    b) σ^2/n
    c) $\sigma^2 * n$
    d) σ/n
28. The probability distribution of the number of successes in a fixed number of independent Bernoulli trials with constant probability of success p is given by the
    a) Poisson distribution
    b) Normal distribution
    c) Exponential distribution
    d) Binomial distribution
29. The sample mean $\bar{X}$ is an unbiased estimator of the population
    a) Median
    b) Mean
    c) Mode
    d) Variance
30. The convergence in mean square of a sequence of random variables Xn to a random variable X means that
    a) The sample mean converges to the population mean
    b) The sample variance converges to the population variance
    c) The sample mean square error approaches zero
    d) The sample distribution approaches a normal distribution
31. Define random sample in the context of statistics.
32. What is the sampling distribution of a statistic?
33. State the formula for the standard error of the sample mean.
34. Define the null hypothesis in hypothesis testing.
35. What is Type I error in hypothesis testing?
36. When is a one-sample proportion test used?
37. What is the critical region in hypothesis testing?
38. How is the p-value approach used in hypothesis testing?
39. State the formula for the standard error of the sample variance.

40. When do we use the two-sample t-test for testing the difference of two means?
41. Define the chi-square distribution with n degrees of freedom.
42. How is the probability density function (p.d.f.) of the chi-square distribution derived using the moment-generating function (m.g.f.)?
43. State the formula for the mean of the chi-square distribution.
44. What is the additive property of the chi-square distribution?
45. Describe the nature of the p.d.f. curve of the chi-square distribution for different degrees of freedom.
46. What are the applications of the chi-square distribution in tests of significance?
47. Define the cumulant generating function in the context of the chi-square distribution.
48. What is the limiting form of the chi-square distribution for a large number of degrees of freedom?
49. How do we construct confidence intervals based on the chi-square distribution?
50. Mention a situation where the chi-square test is commonly used in practice.
51. Define the t-distribution.
52. How is the probability density function (p.d.f.) of the t-distribution derived?
53. State the formula for the mean of the t-distribution.
54. Describe the nature of the probability curve of the t-distribution with different degrees of freedom.
55. What is the relationship between the t, F, and chi-square distributions?
56. Define the F-distribution.
57. How is the p.d.f. of the F-distribution derived?
58. State the formula for the variance of the F-distribution.
59. What is the mode of the F-distribution?
60. How do we perform hypothesis tests based on the t-distribution and F-distribution?
61. Define Parameter and Statistics.
62. What is Standard error?
63. Standard deviation of sample mean is ________.
64. Standard deviation of Sample Proportion is _________.
65. Standard deviation of sample variance is _________.
66. What is Null Hypothesis?
67. When do you suggest two tailed test?
68. Reject H_0 when it is true is ___________ error.
69. _____ is referred as Consumer's risk.
70. P{ Reject a lot when it is good } = ______
71. What is critical region?
72. What is level of significance?
73. What is critical value?
74. What is the test statistic for Single proportion?

TWO MARK QUESTIONS:

1. A random sample of 50 students was taken from a school, and their average score in a math test was 75 with a standard deviation of 5. Calculate the standard error of the sample mean.
2. A manufacturer claims that the proportion of defective products in a batch is 0.10. A sample of 200 products was tested, and 25 were found to be defective. Test the manufacturer's claim using the classical approach.
3. The weight of 10 bags of sugar is recorded as follows (in kg): 5.2, 5.1, 4.9, 5.3, 5.0, 5.2, 4.8, 5.5, 5.4, 5.6. Calculate the sample variance.
4. In a study, the average height of a sample of 100 men was found to be 175 cm with a standard deviation of 6 cm. Test whether the population mean height is significantly different from 180 cm using the p-value approach with a significance level of 0.05.
5. A company claims that their new production method reduces the standard deviation of the product's weight from 2 grams to 1.5 grams. Test this claim using the classical approach with a sample of 50 products. The sample standard deviation is 1.7 grams.
6. Calculate the mean, variance, and standard deviation of the chi-square distribution with 8 degrees of freedom.
7. A research study records the number of cars passing through an intersection during 10 one-hour intervals. The observed and expected frequencies are as follows:
8. Observed: 25, 30, 40, 35, 20, 45, 50, 40, 28, 32 Expected: 30, 35, 45, 38, 22, 40, 52, 39, 25, 30

9. Calculate the chi-square test statistic and test the hypothesis that the observed frequencies follow the expected frequencies.
10. The m.g.f. of a chi-square distribution is given by $M(t) = (1 - 2t)^{(-a/2)}$. Calculate the mean and variance of the chi-square distribution.
11. In an experiment, a six-sided die is rolled 150 times, and the frequencies of each outcome are recorded:
12. Outcome: 1, 2, 3, 4, 5, 6 Frequency: 30, 25, 35, 20, 22, 18
13. Test the hypothesis that the die is fair (equally likely outcomes) using the chi-square test with a significance level of 0.05.
14. The data below shows the number of students studying different subjects in a school:
15. Subjects: Math, Science, English, History Number of Students: 80, 60, 50, 40
16. Calculate the chi-square test statistic to test whether there is a significant difference in the number of students studying different subjects.
17. The t-distribution with 25 degrees of freedom has a mean of 0 and a standard deviation of 1.2. Find the value of t for which $P(T < t) = 0.025$.
18. A researcher collected two samples to compare the average scores of two groups. The sample statistics are as follows:
19. Group 1 ($n_1 = 25$): Mean = 70, Standard Deviation = 8 Group 2 ($n_2 = 30$): Mean = 65, Standard Deviation = 7
20. Test the hypothesis that the two groups have equal population means using the classical approach with a significance level of 0.05.
21. A study recorded the reaction times of individuals under two different conditions. The results are as follows:
22. Condition 1 ($n_1 = 15$): Mean = 0.5 seconds, Standard Deviation = 0.1 seconds Condition 2 ($n_2 = 20$): Mean = 0.6 seconds, Standard Deviation = 0.15 seconds
23. Test the hypothesis that the population standard deviations are equal using the p-value approach with a significance level of 0.01.
24. A random sample of 100 students was taken to compare their scores before and after a training program. The mean difference in scores was 5 with a standard deviation of 3. Test whether the training program has a significant effect on scores using the t-distribution with a significance level of 0.05.
25. The F-distribution with 6 and 8 degrees of freedom has a mean of 1.5 and a variance of 3. Calculate the mode of the F-distribution.
26. Explain the difference between a parameter and a statistic in statistics. Provide an example of each.
27. A sample of 50 students was taken from a population, and the sample mean was found to be 75 with a standard deviation of 10. Determine the 95% confidence interval for the population mean.
28. State the null and alternative hypotheses for each of the following scenarios:
    a) A company claims that their new product has a defect rate of 5%.
    b) A researcher wants to test if the mean IQ score of a sample of students is different from 100.
29. A survey found that 120 out of 400 participants preferred Brand A of a product. Test whether the proportion of people who prefer Brand A is different from 0.30 using the p-value approach with a significance level of 0.05.
30. A sample of 25 observations has a mean of 45 and a standard deviation of 5. Test the hypothesis that the population mean is equal to 50 using the classical approach with a significance level of 0.01.
31. Derive the probability density function (p.d.f.) of the chi-square distribution with 10 degrees of freedom using the moment-generating function (m.g.f.).
32. The following are the observed frequencies of data points in different categories: Category: A, B, C, D, E Frequency: 10, 8, 15, 12, 5
33. Calculate the chi-square test statistic to test whether the observed frequencies follow a uniform distribution.
34. A random variable X follows a chi-square distribution with 12 degrees of freedom. Calculate the probability that X is greater than 20.
35. Prove the additive property of the chi-square distribution, i.e., if $X_1 \sim \chi^2(k_1)$ and $X_2 \sim \chi^2(k_2)$ are independent chi-square variables, then $X_1 + X_2 \sim \chi^2(k_1 + k_2)$.
36. A study on the preferences of soft drinks categorized the responses into three options: Coke, Pepsi, and Other. The observed frequencies are as follows
    Coke: 120, Pepsi: 80, Other: 50

37. Perform a chi-square goodness-of-fit test to determine if the distribution of preferences is different from what would be expected under equal preference.
38. Derive the probability density function (p.d.f.) of the t-distribution with 15 degrees of freedom.
39. Calculate the value of t for which $P(T > t) = 0.025$ for a t-distribution with 30 degrees of freedom.
40. A random sample of 20 observations has a mean of 65 and a standard deviation of 8. Test the hypothesis that the population mean is equal to 70 using the t-distribution with a significance level of 0.05.
41. The heights of students from two schools, A and B, were recorded. The sample statistics are as follows:
42. School A ($n1 = 30$): Mean = 160 cm, Standard Deviation = 5 cm School B ($n2 = 25$): Mean = 165 cm, Standard Deviation = 6 cm
43. Test the hypothesis that the population means of the two schools are equal using the t-distribution with a significance level of 0.01.
44. Derive the probability density function (p.d.f.) of the F-distribution with 5 and 10 degrees of freedom.
45. The following are the observed variances of two independent samples: Sample 1 ($n1 = 20$): Variance = 16 Sample 2 ($n2 = 15$): Variance = 12
46. Test the hypothesis that the population variances are equal using the F-distribution with a significance level of 0.05.
47. Standard error of difference of two sample proportions is __________.
48. The standard deviation of a population is 5. If a sample if size 100 is drawn, what is the standard error of the sample standard deviation ?
49. What is test statistics for Difference of mean.

LONG TYPE QUESTIONS:

1. A company is testing a new manufacturing process for producing light bulbs. A random sample of 100 light bulbs is taken, and their lifetimes are recorded. The sample mean lifetime is 800 hours with a standard deviation of 50 hours. Test whether the new manufacturing process has significantly increased the mean lifetime of light bulbs compared to the previous process, using the p-value approach with a significance level of 0.05.
2. A political survey was conducted to determine the proportion of voters in favor of a particular candidate. The survey sampled 1200 voters, of which 550 were in favor of the candidate. Test whether there is a significant difference between the proportions of male and female voters in favor of the candidate, using the classical approach with a significance level of 0.01.
3. A quality control engineer wants to test whether two production lines have the same variance in their product weights. He takes a sample of 50 products from each line and records their weights. The sample standard deviations are 4 grams and 3.5 grams for Line 1 and Line 2, respectively. Perform an F-test to test the null hypothesis that the population variances are equal, with a significance level of 0.05.
4. A researcher wants to compare the average scores of three different groups of students on a standardized test. The sample statistics for each group are as follows:
   Group 1 ($n1 = 30$)
   Mean = 75, Standard Deviation = 10
   Group 2 ($n2 = 25$)
   Mean = 70, Standard Deviation = 12
   Group 3 ($n3 = 35$)
   Mean = 80, Standard Deviation = 8
   Perform an analysis of variance (ANOVA) to determine if there is a significant difference in the mean scores of the three groups, with a significance level of 0.01.
5. A researcher is conducting a hypothesis test to compare the mean income of two different cities. The sample statistics are as follows:
   City A ($n1 = 100$): Mean = \$45,000, Standard Deviation = \$5,000
   City B ($n2 = 120$): Mean = \$50,000, Standard Deviation = \$4,000
   Perform a two-sample t-test to test whether the mean income of City B is significantly higher than City A, with a significance level of 0.05.

6. Prove that the chi-square distribution is the limiting form of the gamma distribution as the number of degrees of freedom approaches infinity.
7. A study on the preferences for different ice cream flavors collected data from three different age groups: Young, Middle-aged, and Elderly. The observed frequencies and expected frequencies (under the assumption of equal preferences) are as follows:
   Young: Observed = 100, Expected = 80
   Middle-aged: Observed = 80, Expected = 80
   Elderly: Observed = 60, Expected = 80
   Perform a chi-square goodness-of-fit test to determine if there is a significant difference in ice cream flavor preferences across age groups, with a significance level of 0.05.
8. A company claims that its product has a failure rate of 2%. To test this claim, a sample of 500 products is taken, and 14 of them are found to be defective. Test the company's claim using the chi-square test for goodness of fit, with a significance level of 0.01.
9. A research study compares the performance of three different teaching methods on a group of students. The data collected is as follows:
   Teaching Method 1: 80, 85, 75, 90, 88
   Teaching Method 2: 92, 95, 88, 90, 85, 88, 82
   Teaching Method 3: 78, 82, 80, 85, 90, 86
   Perform a one-way ANOVA to test whether there is a significant difference in the mean performance among the three teaching methods, with a significance level of 0.05.
10. A medical study is conducted to determine the effectiveness of three different treatments for a certain condition. The number of patients showing improvement and no improvement after each treatment is recorded. The data is as follows:
    Treatment 1: Improved = 30, Not Improved = 20
    Treatment 2: Improved = 40, Not Improved = 10
    Treatment 3: Improved = 25, Not Improved = 15
    Perform a chi-square test for independence to determine if there is a significant association between the treatment and improvement status, with a significance level of 0.01.
11. Derive the probability density function (p.d.f.) of the t-distribution with 30 degrees of freedom.
12. A researcher wants to determine if there is a significant difference in the mean test scores of students from three different schools. The sample statistics for each school are as follows:
    School 1 (n1 = 40): Mean = 80, Standard Deviation = 10
    School 2 (n2 = 35): Mean = 75, Standard Deviation = 12
    School 3 (n3 = 30): Mean = 85, Standard Deviation = 8
    Perform an analysis of variance (ANOVA) to test whether there is a significant difference in the mean scores of the three schools, with a significance level of 0.05.
13. A study on the effect of two different diets on weight loss recorded the following data:
    Diet 1 (n1 = 25): Mean weight loss = 6 kg, Standard Deviation = 1 kg
    Diet 2 (n2 = 30): Mean weight loss = 7 kg, Standard Deviation = 1.5 kg
    Perform a two-sample t-test to test whether there is a significant difference in weight loss between the two diets, with a significance level of 0.01.
14. The lifetimes (in hours) of two different brands of batteries were recorded. The sample statistics for each brand are as follows:
    Brand A (n1 = 50): Mean = 500, Standard Deviation = 50
    Brand B (n2 = 60): Mean = 480, Standard Deviation = 55
    Perform a two-sample t-test to determine if there is a significant difference in the mean lifetimes of the two brands, with a significance level of 0.05.
15. A study on the effects of two different training programs on athletes' performance recorded the following data:
    Program 1 (n1 = 20): Mean performance = 8.2 seconds, Standard Deviation = 0.5 seconds Program 2 (n2 = 25): Mean performance = 7.8 seconds, Standard Deviation = 0.4 seconds
    Perform a two-sample t-test to test whether there is a significant difference in performance improvement between the two programs, with a significance level of 0.05.

16. Explain the concept of a sampling distribution of a statistic. How is it different from the sampling distribution of a sample mean?
17. Discuss the significance of the null and alternative hypotheses in hypothesis testing. Explain how Type I and Type II errors are related to these hypotheses.
18. Define the level of significance in hypothesis testing. How does it influence the decision-making process in hypothesis testing?
19. Compare and contrast the classical approach and the p-value approach in hypothesis testing. When would you prefer one approach over the other?
20. Discuss the conditions that must be satisfied to perform a large sample test for a single proportion. State the formula for the test statistic and the critical region for a given significance level.
21. Derive the probability density function (p.d.f.) of the chi-square distribution with n degrees of freedom using the moment-generating function (m.g.f.). Discuss the properties of the chi-square distribution curve for different values of n.
22. Define the cumulative generating function and its role in the chi-square distribution. How does it help in calculating moments of the distribution?
23. Describe the applications of the chi-square distribution in tests of significance and constructing confidence intervals. Explain why it is commonly used in various statistical analyses.
24. Prove the relationship between the chi-square distribution and the normal distribution in the context of the central limit theorem. How does this relationship facilitate hypothesis testing and interval estimation?
25. Discuss the difference between exact and approximate sampling distributions. Provide examples of statistical tests and confidence intervals based on the chi-square distribution.
26. Derive the probability density function (p.d.f.) of the t-distribution and the F-distribution. Explain the role of degrees of freedom in determining the shape of these distributions.
27. Discuss the nature of the t-distribution and the F-distribution curves for different degrees of freedom. How does the sample size affect the shape of these distributions?
28. Prove the relationship between the t-distribution and the standard normal distribution in the context of small sample size and large sample size.
29. Explain the application of the t-distribution in hypothesis testing and constructing confidence intervals for a single mean and the difference between two means.