DO NOT OPEN THIS TEST BOOKLET UNTIL YOU ARE ASKED TO DO SO

TEST BOOKLET

Sl. No. 0256

Subject Code : 26

Subject : Statistics

LECTURERS FOR NON-GOVT. AIDED COLLEGES OF ODISHA

Time Allowed : 2 Hours

Maximum Marks : 150

: INSTRUCTIONS TO CANDIDATES :

1. **IMMEDIATELY AFTER THE COMMENCEMENT OF THE EXAMINATION, YOU SHOULD CHECK THAT THIS TEST BOOKLET CONTAINS 24 PAGES AND DOES NOT HAVE ANY UNPRINTED OR TORN OR MISSING PAGES OR ITEMS ETC. IF SO, GET IT REPLACED BY A COMPLETE TEST BOOKLET.**

2. You have to enter your **Roll No.** on the Test Booklet in the Box provided alongside. **DO NOT** write anything else on the Test Booklet.

3. The Test Booklet contains **100** questions. Each question comprises four answers. You have to select the correct answer which you want to mark (darken) on the **Answer Sheet (OMR Sheet)**. In any case choose ONLY ONE answer for each question. If more than one answer is darkened, it will be considered as wrong.

4. You have to mark (darken) all your answers only on the **OMR Answer Sheet using BLACK BALL POINT PEN** provided by the State Selection Board. You have to do rough work only in the space provided at the end of the Test Booklet. See instructions in the Answer Sheet.

5. All questions carry equal marks i.e. of one and half mark for each correct answer and each wrong answer will result in negative marking of **0.50** mark.

6. Before you proceed to mark (darken) the answers in the **OMR Answer Sheet** to the questions in the Test Booklet, you have to fill in some particulars in the Answer Sheet as per the instructions in your Admit Card.

7. On completion of the examination, you should hand over the **original Answer Sheet (OMR Sheet)** issued to you to the Invigilator before leaving the Examination Hall. You are allowed to take with you the candidate's copy (carbon copy) of the **OMR Answer Sheet** along with the Test Booklet for your reference.

Candidate's full signature

Invigilator's signature

SEAL

1. The sum of two positive quantities is 2n. Find the chance that the product of the two quantities is not less than 3/4 times their greatest product :

(A) 0.30

(B) 0.50

(C) 0.25

(D) 0.45

2. The probability that a teacher will give an unannounced test during any class meeting is 0.2. If a student is absent twice, what is the probability that he will miss at least one test ?

(A) 0.64

(B) 0.80

(C) 0.60

(D) 0.36

3. Company A produces 10% defective products, company B produces 20% defective products and company C produces 5% defective products. If choosing a company is an equally likely event, then find the probability that the product chosen is defective :

(A) 0.22

(B) 0.12

(C) 0.10

(D) 0.21

4. A survey determines that in a locality, 33% go to work by bike, 42% go by car, and 12% use both. The probability that a random person selected uses neither of them is :

(A) 0.39

(B) 0.37

(C) 0.63

(D) 0.75

5. At a certain university, 4% of men are over 6 feet tall and 1% of women are over 6 feet tall. The total student population is divided in the ratio 3 : 2 in favour of women. If a student is selected at random from among all those over six feet tall, what is the probability that the student is a woman ?

(A) $\frac{8}{11}$

(B) $\frac{3}{5}$

(C) $\frac{3}{11}$

(D) $\frac{1}{20}$

6. For some constant c, the random variable X has probability density function :

$$f(x) = \begin{cases} cx^n, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $P(X > x)$.

(A) $x^{n+1} - 1$

(B) x^{n+1}

(C) 1

(D) $1 - x^{n+1}$

7. If a random variable X has the density function

$$f(x) = \begin{cases} \frac{1}{2}e^{-|x|}, & -\infty < x < \infty \\ 0, & \text{otherwise} \end{cases}$$

the distribution function is given by :

(A) $F(x) = \begin{cases} \frac{e^x}{2}, x \le 0 \\ 1 - \frac{e^x}{2}, x > 0 \end{cases}$

(B) $F(x) = \begin{cases} \frac{e^x}{2}, x \le 0 \\ 1 - \frac{e^{-x}}{2}, x > 0 \end{cases}$

(C) $F(x) = \begin{cases} 1 - \frac{e^{-x}}{2}, 0 < x < \infty \\ 1, \text{ otherwise} \end{cases}$

(D) $F(x) = \begin{cases} 1 - \frac{e^x}{2}, x < 0 \\ \frac{e^{-x}}{2}, x \ge 0 \end{cases}$

8. Pick out the correct relation among different modes of convergence of a sequence of random variables :

(A) Convergence almost surely $\Rightarrow$ convergence in distribution $\Rightarrow$ convergence in probability

(B) Convergence in r-th mean $\Rightarrow$ convergence almost surely $\Rightarrow$ convergence in probability

(C) Convergence in probability $\Rightarrow$ convergence in r-th mean $\Rightarrow$ convergence in distribution

(D) Convergence in r-th mean $\Rightarrow$ convergence in probability $\Rightarrow$ convergence in distribution

9. Let $\{X_n\}$ be a sequence of independent random variables with $E(X_k) = \mu_k$ and $V(X_k) = \sigma_k^2$. Then,

$$\lim_{n \to \infty} \frac{1}{n^2} \Sigma_{k=1}^{n} \sigma_k^2 \to 0 \Rightarrow$$

$$\frac{\Sigma_{k=1}^{n} X_k - \Sigma_{k=1}^{n} \mu_k}{n} \xrightarrow{P} 0$$

This theorem is called :

(A) Bernoulli's theorem on WLLN

(B) Khinchin's theorem on WLLN

(C) Chebyshev's theorem on WLLN

(D) Poisson's theorem on WLLN

10. Lindeberg-Levy Central Limit Theorem assumes that the sequence of random variables under consideration are :

(A) i.i.d. with a common mean and a common variance ($< \infty$)

(B) i.i.d. Bernoulli variables with common parameter p

(C) independent with different means and different variances ($< \infty$)

(D) independent and their third and higher order moments exist

11. Find the moment generating function of X if its probability density function is

$$f(x) = \begin{cases} x \text{ for } 0 < x < 1 \\ 2 - x \text{ for } 1 < x < 2 \end{cases} :$$

(A) $\left(\frac{e^{-t} - 1}{t}\right)^2$

(B) $\left(\frac{e^{t} - 1}{t}\right)^2$

(C) $\left(\frac{e^{2t} - 1}{t}\right)^2$

(D) $\frac{e^{2t} - 1}{t^2}$

12. Let X and Y be jointly distributed continuous random variables with joint probability density function

$$f(x, y) = \begin{cases} 6e^{-(2x+3y)}, & x, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Find $E(Y|X > 2)$:

(A) $\frac{1}{3}$

(B) $\frac{2}{3}$

(C) $\frac{1}{6}$

(D) $\frac{3}{5}$

13. If a binomial variate has mean 4 and variance 3, find its third central moment μ_3 :

(A) $\frac{1}{2}$

(B) $\frac{5}{2}$

(C) $\frac{3}{2}$

(D) $\frac{3}{4}$

14. If $X \sim N(\mu, \sigma^2)$, find variance of $\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2$:

(A) $\frac{1}{4}$

(B) $\frac{1}{2}$

(C) 1

(D) $\frac{1}{8}$

15. If a random variable X has a beta distribution of the first kind with parameters $m > 1$ and $n > 1$, the mode lies at the point :

(A) $\frac{m}{m+n-2}$

(B) $\frac{m-1}{m+n-1}$

(C) $\frac{m-1}{m+n-2}$

(D) $\frac{m}{m+n}$

16. The life time (X) in hours of a certain electrical equipment has a normal distribution with mean 80. However, experience shows that life times of approximately 99% of the said equipment lie in the interval [68, 92]. Find variance of X :

(A) 25

(B) 36

(C) 144

(D) 16

17. Which one is not a condition of a Poisson model with probability of having success p ?

(A) p in a small time interval is constant

(B) p more than one in a small time interval is very small

(C) p in a small time interval is independent of time and also of earlier success

(D) p in a small time interval (t, t + dt) is kdt for a positive constant k

18. Buses arrive at a specified stop at 15 minutes intervals starting at 7 AM. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7 : 30, find the probability that he waits less than 5 minutes for a bus :

(A) $\frac{1}{6}$

(B) $\frac{2}{3}$

(C) $\frac{1}{3}$

(D) $\frac{3}{4}$

19. Find the variance of the number of times one must throw a die until the outcome 1 has occurred 4 times :

(A) 144

(B) 120

(C) 124

(D) 96

20. If X and Y have a bivariate normal distribution with means $\mu_X = 12$, $\mu_Y = 10$; standard deviations $\sigma_X = 30$, $\sigma_Y = 20$ and correlation coefficient $\rho = 0.8$, find $E(X|Y = 9)$:

(A) 12.00

(B) 10.80

(C) 13.20

(D) 12.53

21. For the frequency distribution with classes of unequal widths, the heights of bars of the histogram are proportional to :

(A) Frequencies in percentage

(B) Class intervals

(C) Frequency densities

(D) Class frequencies

22. The suitable graph to represent the time series data is :

(A) Historigram

(B) Histogram

(C) Frequency curve

(D) Frequency polygon

23. _____ curve is used to study disparities of the distributions of income and wealth for certain segment of a population.

(A) Ogive curve

(B) Frequency curve

(C) Range curve

(D) Lorenz curve

24. 120 students appeared for a certain test and the following marks distribution was obtained. If 35 marks are required for passing, find an approximate number of failed students :

Marks	Students
0 – 20	10
20 – 40	30
40 – 60	36
60 – 80	30
80 – 100	14

(A) 33

(B) 17

(C) 41

(D) 35

25. Karl Pearson's coefficient of skewness of a distribution is +0.40 whose mean and standard deviations are 30 and 8 respectively. Find the mode of the distribution :

(A) 28.9

(B) 26.8

(C) 21.8

(D) 18.5

26. The first three moments of a distribution about value 1 are 2, 25 and 80 respectively. Find variance and μ_3 :

(A) 21 and – 54

(B) 21 and – 30

(C) 16 and – 71

(D) 25 and – 44

27. Arithmetic and harmonic means of x_i, $i = 1, 2, ..., n$, are $\bar{x}$ and h respectively. Then the harmonic mean H_u of $u_i = \frac{1}{\delta x_i}$, $i = 1, 2, ..., n$, is given by :

(A) $H_u = \frac{1}{\delta h}$

(B) $H_u = \frac{1}{\delta}\bar{x}$

(C) $H_u = \frac{1}{\delta \bar{x}}$

(D) $H_u = \frac{\delta}{h}$

28. The first normal equation, for fitting a curve of the type $y = ab^x$ to a data set (x_i, y_i), $i = 1, 2, ..., n$ by the least squares method, is $\Sigma \log y = n \log a + \log b\, \Sigma x$. The second equation is :

(A) $\Sigma x \log y = \log a\, \Sigma x + \log b\, \Sigma x^2$

(B) $\Sigma x \log y = \log a\, \Sigma x + \log b\, (\Sigma x)^2$

(C) $\Sigma x^2 \log y = \log a\, \Sigma x + \log b\, (\Sigma x)^2$

(D) $\Sigma x^2 \log y = \log a\, \Sigma x \log y + \log b\, \Sigma x\, \Sigma \log y$

29. In a very hotly fought battle, at least 70% of the combatants lost an eye, at least 75% lost an year, at least 80% an arm and at least 85% a leg. Find at least what percentage of combatants lost all four :

(A) 17%

(B) 19%

(C) 15%

(D) 10%

30. If X and Y are two correlated variables having the same standard deviation σ and the correlation coefficient ρ, then the correlation coefficient between X and X + Y is :

(A) $\frac{1}{\sqrt{1+\rho^2}}$

(B) $\frac{\sigma}{\sqrt{2(1+\rho)}}$

(C) $\sqrt{\frac{1+\rho}{2}}$

(D) $\frac{\sqrt{1+\rho}}{2\sigma}$

31. The two regression equations of Y on X and X on Y are $Y = \theta X + 4$ and $X = 4Y + 5$ respectively. A plausible interval for θ is :

(A) (0, 0.25)

(B) (0, 0.50)

(C) (0. 20,1.0)

(D) (0, 1.0)

32. The coefficient of rank correlation of the marks obtained by 10 students in two subjects was 0.4. But, later it was found that the difference in ranks for one student was wrongly taken as 2 instead of 3. Find the correct coefficient of rank correlation :

(A) 0.43

(B) 0.37

(C) 0.52

(D) 0.25

33. For a trivariate distribution in X_1, X_2 and X_3, all the total correlation coefficient between the variables are equal, say α. If $R_{1.23}$ is the multiple correlation coefficient of X_1 on X_2 and X_3, then $1 - R^2_{1.23}$ is equal to :

(A) $(1-\alpha)(1+2\alpha)/(1+\alpha)$

(B) $[(1+\alpha)-(1-\alpha)(1+2\alpha)]/(1+\alpha)$

(C) $(1-\alpha)(1-2\alpha)/(1+\alpha)$

(D) $(1+\alpha)(1-2\alpha)/(1-\alpha)$

34. If ρ and η are respectively the correlation coefficient and correlation ratio while dealing with two variables X and Y, then which one of the following statement is not true ?

(A) η is an appropriate measure of curvilinear relationship between the variables

(B) When both variables are dichotomous, η is more suitable than ρ to measure association between them

(C) $1-\eta^2 \leq 1-\rho^2$

(D) The value of η is not independent of the classification of data on the variables

35. When a random sample $x_1, x_2, ..., x_n$ is taken from a normal population $N(\mu, \sigma^2)$, the variance of the sample variance $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is :

(A) $2\sigma^2/n$

(B) $\sigma^2/2n$

(C) $2\sigma^4/n$

(D) σ^4/n^2

36. If X and Y are independent χ^2 variates with 6 and 8 degrees of freedom, then X/Y is a :

(A) $\beta_1(3, 4)$ variate

(B) χ^2 variate with 14 degrees of freedom

(C) F(5, 7) variate

(D) $\beta_2(3, 4)$ variate

37. For $4 < n < 30$, the t-distribution curve with regard to flatness is :

(A) Mesokurtic

(B) Bimodal

(C) Leptokurtic

(D) Platykurtic

38. The modal value of a F-distribution with $(8, \nu_2)$ degres of freedom is 7/12. Compute ν_2.

(A) 5

(B) 7

(C) 11

(D) 10

39. If χ^2_n is a chi-square variate with n degrees of freedom, then $\sqrt{\chi^2_n}$ will be distributed as a :

(A) Chi-square distribution

(B) Exponential distribution

(C) Gamma distribution

(D) Fisher's t-distribution

40. If X has a F-distribution with (m, n) degrees of freedom, then the distribution of 1/X will be :

(A) t-distribution with n degree of freedom

(B) F-distribution with $\left(\frac{n}{2}, \frac{m}{2}\right)$ degrees of freedom

(C) F-distribution with (n, m) degrees of freedom

(D) Chi-square distribution with m degree of freedom

41. _____ property ensures that the difference between the estimator T_n and the estimable parametric function $\gamma(\theta)$ decreases as the sample size n increases indefinitely.

(A) Sufficiency

(B) Consistency

(C) Unbiasedness

(D) Efficiency

42. If $\bar{x}$ is the sample mean based on random sample of size n drawn from the exponential distribution $f(x) = \theta e^{-\theta x}$, $0 \leq x < \infty$, $\theta > 0$, an estimator of the parameter θ obtained by the method of moments is :

(A) $1/\bar{x}$

(B) $1/\bar{x}^2$

(C) $\bar{x}$

(D) $\bar{x}^2$

43. Which one of the following statements about a maximum likelihood estimator, if exists, is not true ?

(A) It is consistent

(B) It is a function of the sufficient statistic

(C) It is asymptotically normally distributed about the true parameter value

(D) It is uniformly minimum variance unbiased

44. Which of the following is the most commonly used estimation tool in regression analysis ?

(A) Method of moments

(B) Method of minimum chi-square

(C) Method of least squares

(D) Method of maximum likelihood

45. Let U be the set of all unbiased estimators of $\gamma(\theta)$, $\forall\ \theta \in \Theta$. An estimator $T_0 \in U$ is such that $V_\theta(T_0) \leq V_\theta(T)$ for all $\theta \in \Theta$ and for every $T \in U$. Then T_0 is called :

(A) MVB estimator of $\gamma(\theta)$

(B) UMVU estimator of $\gamma(\theta)$

(C) The most efficient estimator of $\gamma(\theta)$

(D) Minimal sufficient estimator of $\gamma(\theta)$

46. For a Cauchy's distribution whose probability density function is given by $f(x) = \frac{1}{\pi\left[1+(x-\theta)^2\right]}, -\infty < x < \infty$:

(A) The sample median is an unbiased and consistent estimator of the population mean

(B) The sample mean is an unbiased and consistent estimator of the population mean

(C) The sample mean is a biased but consistent estimator of the population mean

(D) The sample median is an unbiased but not consistent estimator of the population mean

47. If a random sample $x_1, x_2, \ldots, x_n$ is taken from a normal population $N(0, \sigma^2)$, the Cramer-Rao lower bound for estimating σ^2 is :

(A) $2\sigma^2 / n$

(B) $3\sigma^2 / n^2$

(C) $\sigma^4 / 2n$

(D) $2\sigma^4 / n$

48. If $\bar{x}$ is the sample mean based on a random sample of n observations from a binomial population B(N, p), the maximum likelihood estimator of p is :

(A) $\bar{x}$

(B) $\bar{x} / N$

(C) $n\bar{x} / N$

(D) $\bar{x} / n$

49. Consider a random variable X with probability density function f(x). Then for which of the following there does not exist a sufficient statistic for θ ?

(A) $f(x) = \frac{1}{\pi\left[1+(x-\theta)^2\right]}, -\infty < x < \infty$

(B) $f(x) = \frac{x}{\theta}, 0 \le x < \theta$

(C) $f(x) = \frac{1}{2\pi} e^{-\frac{1}{2}(x-\theta)^2}, -\infty < x < \infty$

(D) $f(x) = \frac{1}{\theta} e^{-x/\theta}, 0 < x < \infty$

50. A student conducted a study and reported that the 95% confidence interval for the population mean was [46, 52]. He was sure that the population SD was 16. What was the sample size (rounded to the nearest whole number) used to compute the said interval ?

(A) 62

(B) 97

(C) 109

(D) 40

51. The p-value of a test is the :

(A) Largest significance level for which the null hypothesis can be rejected

(B) Smallest significance level at which the null hypothesis can be rejected

(C) Smallest significance level at which the null hypothesis can be accepted

(D) Probability that no errors have been made in rejecting or accepting the null hypothesis

52. When a coin is thrown independently 4 times to test the hypothesis that the probability of heads is $\frac{1}{2}$ versus the alternative that the probability is $\frac{2}{3}$, the size of the type – 1 error is :

(A) $\frac{5}{16}$

(B) $\frac{16}{27}$

(C) $\frac{11}{16}$

(D) $\frac{11}{27}$

53. If a random sample $x_1, x_2, ..., x_n$ is taken from a normal population $N(\theta, 1)$, then the best critical region of size α under the Neyman-Pearson paradigm to test the hypothesis $H_0 : \theta = 0$ against $H_1 : \theta = 1$ is given by :

(A) $\frac{n}{2} - \log k_\alpha < \Sigma_{i=1}^{n} x_i < \frac{n}{2} + \log k_\alpha$

(B) $\Sigma_{i=1}^{n} x_i < \frac{n}{2} + \log k_\alpha$

(C) $\Sigma_{i=1}^{n} x_i < \frac{n}{2} - \log k_\alpha$

(D) $\Sigma_{i=1}^{n} x_i > \frac{n}{2} + \log k_\alpha$

54. A test based on a critical region W of size α for testing $H_0 : \theta = \theta_0$ against a composite alternative $H_1 : \theta \neq \theta_0$ is such that $P\{x \in W/H_1\} \geq P\{x \in W_1/H_1\}$ for all $\theta \neq \theta_0$ and for every critical region W_1 of size α. Then the test is called :

(A) A most powerful test

(B) An uniformly most powerful test

(C) An unbiased test

(D) A most powerful unbiased test

55. Point out the incorrect statement regarding the likelihood ratio test :

(A) Likelihood ratio is a random variable lying between 0 and 1.

(B) The test is generally UMP if an UMP test exists.

(C) Under certain conditions, the likelihood ratio is normally distributed.

(D) The test is asymptotically consistent but not necessarily unbiased.

56. Median test utilizes :

(A) Hypergeometric distribution

(B) Negative binomial distribution

(C) Binomial distribution

(D) Poisson distribution

57. If the sample sizes n_1 and n_2 of the two samples are large, the distribution of Mann-Whitney U statistic is asymptotically normal with variance :

(A) $\frac{n_1 n_2 (n_1 + n_2)}{12}$

(B) $\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$

(C) $\frac{n_1 n_2 (n_1 n_2 + 1)}{12}$

(D) $\frac{n_1 n_2 (n_1 + n_2 - 1)}{12}$

58. It is thought that 20% of all students taking a particular course received a grade of A. In a sample of 100 students, it is found that 15 made an A. What is the calculated value of the test statistic for testing the true proportion is 20% under large sample theory ?

(A) 1.40

(B) 3.12

(C) 2.35

(D) 1.25

59. Genetic theory states that the proportion of children of the parents of three different blood types M, MN and N on the average will be 1 : 2 : 1. But a survey report states that in a random sample of 300 children, the numbers of children belonging to the types M, MN and N were respectively 90, 135 and 75. Value of chi-square statistic for 2 degrees of freedom calculated from the data in order to test the correspondence between the theoretical and observed frequencies is :

(A) 4.50

(B) 4.17

(C) 11.25

(D) 9.03

60. It is known that for right-handed people, the dominant (right) hand tends to be stronger. For left-handed people who live in a world designed for right-handed people, the same may not be true. To test this, muscle strength was measured on the right and left hands of a random sample of 15 left-handed men and the difference (left-right) was found. The alternative hypothesis is one-sided (left hand stronger). The resulting t-statistic was 1.80. This is an example of :

(A) A two-sample t-test

(B) A paired t-test

(C) A pooled t-test

(D) A one-sample t-test

61. Which of the following statements is incorrect for a sample survey ?

(A) A sample survey is a time saving and cost reducing method

(B) A sample survey is subject to both sampling and non-sampling errors

(C) Sampling errors can be completely eliminated from a sample survey

(D) Non-sampling errors in a sample survey can be controlled to a greater extent than in case of a census survey

62. When a sample of 10 units is drawn from a population of 110 units both by simple random sampling without and with replacements, what is the percentage relative gain in efficiency of the former over the later for estimating population mean ?

(A) 11.2%

(B) 8.2%

(C) 9.0%

(D) 10.0%

63. For a stratified random sampling, let N_j, S_j^2 and C_j respectively be size, variance and cost of the survey per unit in the jth stratum. Then the optimum sample size for j^{th} stratum that minimizes variance of the estimate for a specified total cost is :

(A) $n_j = n \dfrac{N_j S_j / \sqrt{C_j}}{\Sigma_j N_j S_j / \sqrt{C_j}}$

(B) $n_j = n \dfrac{N_j S_j^2 / \sqrt{C_j}}{\Sigma_j N_j S_j^2 / \sqrt{C_j}}$

(C) $n_j = n \dfrac{N_j S_j C_j}{\Sigma_j N_j S_j C_j}$

(D) $n_j = n \dfrac{N_j S_j^2 \sqrt{C_j}}{\Sigma_j N_j S_j^2 \sqrt{C_j}}$

64. While talking about the merits and demerits of multi-stage sampling, which one of the following statements is not true ?

(A) It reduces the costs and time associated with data collection

(B) Provides flexibility by allowing different sampling methods at different stages

(C) The sample will not be 100% representative of the entire population

(D) Normally more accurate than simple random sampling with the same sample size

65. For estimating P, the proportion of units in a population of size N possessing an attribute, a simple random sample of n units is drawn by without replacement. Then 95% confidence limits for P under normal approximation using unbiased estimator p (sample proportion) are :

(A) $p \pm 1.96 \sqrt{\frac{N-n}{(n-1)N} p(1-p)}$

(B) $p \pm \left[1.96 \sqrt{\frac{N-n}{(n-1)N} p(1-p)} + \frac{1}{2n}\right]$

(C) $p \pm 1.96 \sqrt{\frac{N-n}{nN} p(1-p)}$

(D) $p \pm \left[1.96 \sqrt{\frac{N-n}{nN} p(1-p)} + \frac{1}{n}\right]$

66. A random sample of 10 units is drawn from a population of 100 units by simple random sampling with replacement in order to estimate the ratio = $\bar{Y} / \bar{X}$, where $\bar{Y}$ and $\bar{X}$ are population means of y and x. If it is known that $\bar{Y} = 10$, $\bar{X} = 5$ and $\delta = \Sigma_{i=1}^{100} (y_i - Rx_i)^2 = 500$, find variance of the ratio estimator :

(A) 0.18

(B) 0.20

(C) 0.02

(D) 0.03

67. When the number of units in the population is not an integral multiple of sample size, the systematic sampling procedure adopted is called :

(A) Circular systematic sampling

(B) Linear systematic sampling

(C) Two-dimensional systematic sampling

(D) Balanced systematic sampling

68. Suppose that a random sample of n units is drawn from a finite population of N units with unequal probabilities without replacement with y_i as the measurement for the ith unit. Let π_i be inclusion probability of ith unit in the sample, and π_{ij} be the joint inclusion probability of ith and jth units in the sample. For estimating population total which of the following is not a correct expression for the variance of the Horvitz-Thompson estimator ?

(A) $\frac{1}{2}\Sigma_{i \neq j}^{N}\left(\pi_i\pi_j - \pi_{ij}\right)\left(\frac{y_i}{\pi_i} - \frac{y_i}{\pi_j}\right)^2$

(B) $N^2 \Sigma_{i=1}^{N}\Sigma_{j>i}^{N}\left(\pi_i\pi_j - \pi_{ij}\right)\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2$

(C) $\Sigma_{i=1}^{N}\Sigma_{j>i}^{N}\left(\pi_i\pi_j - \pi_{ij}\right)\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2$

(D) $\Sigma_{i=1}^{N}\frac{1-\pi_i}{\pi_i}y_i^2 + 2\Sigma_{i=1}^{N}\Sigma_{j>i}^{N}\frac{\left(\pi_{ij}-\pi_i\pi_j\right)}{\pi_i\pi_j}y_iy_j$

69. In a crop survey it is desired to estimate total area under mustard in a block by regression method of estimation considering village as the sampling unit and cultivated area of the village as an auxilliary variable. If the total cultivated area under the block is not known, what would be the appropriate sampling design to be used ?

(A) Conventional two-stage sampling

(B) Stratified two-phase sampling

(C) Conventional two-phase sampling

(D) Two-phase cluster sampling

70. If a sample of 2 units is drawn from a population of N units by PPS sampling without replacement with P_i as the probability of selecting ith unit at the first draw, then the probability of selecting ith unit at the second draw is :

(A) $P_iP_j\left(\frac{1}{1-P_i} + \frac{1}{1-P_j}\right)$

(B) $P_i + \Sigma_{j \neq i}\frac{P_iP_j}{1-P_j}$

(C) $\Sigma_{j \neq i}\frac{P_iP_j}{1-P_j}$

(D) $\frac{P_i}{1-P_i} + \frac{P_j}{1-P_j}$

71. Local control helps to :

(A) Obtain a valid estimate of the error variance

(B) Reduce the experimental error

(C) Make the experiment free from any systematic influence of environment

(D) Enable to infer whether the differences in the treatment means are actually more than the sampling fluctuations

72. For a resolvable BIBD with parameters ϑ, b, r, k and λ :

(A) $b \geq \vartheta + r - k$

(B) $b \leq \vartheta + k - r$

(C) $b \geq \vartheta + k - 1$

(D) $b \geq \vartheta + r - 1$

73. Identify the incorrect statement :

(A) Split-plot is the appropriate design when two sets of treatments are involved in the experiment requiring large size plots.

(B) A split-plot design is a confounded design in which the main plot treatments are confounded.

(C) In a split-plot design, increased precision is attained on subplot treatments and interaction.

(D) In a split-plot design, the precision of whole plot treatments is less as compared to their precision if they were tried in a randomised block design.

74. Which of the following statements is wrong in respect of CRD ?

(A) It is recommended in situations where some units are likely to be destroyed or fail to respond.

(B) It provides the maximum number of degrees of freedom for estimating experimental error.

(C) It is suitable only for a large number of treatments.

(D) It is completely flexible, i.e., any number of treatments and any number of units per treatment may be used.

75. If the total response values for treatment combinations a_0b_0, a_1b_0, a_0b_1 and a_1b_1 in a factorial experiment with 2 factors A and B each at 2 levels from 4 replications are respectively 18, 17, 25 and 30, find the sum of squares for the interaction effect AB :

(A) 1.35

(B) 2.25

(C) 2.50

(D) 1.00

76. Consider a RBD for t treatments and r blocks with S_B^2 and S_E^2 as the mean squares due to block and error. Then the efficiency of this design relative to CRD, if conducted for the same experiment, is :

(A) $\dfrac{r(t-1)S_E^2 + (r-1)S_B^2}{(rt-1)S_E^2}$

(B) $1 + \dfrac{r(t-1)S_B^2}{(rt-1)S_E^2}$

(C) $\dfrac{(r-1)(t-1)S_E^2 + (r-1)\,S_B^2}{(rt-1)\,S_E^2}$

(D) $\dfrac{(rt-1)S_E^2 + (r-1)\,S_B^2}{(rt-1)\,S_E^2}$

77. If interaction AC is confounded in a 2^3 factorial experiment with factors A, B and C, the arrangement of the treatment combinations for two blocks in a replicate will be :

(A) Block 1 : b, ac, bc, a
Block 2 : (1), ab, c, abc

(B) Block 1 : b, c, ab, ac
Block 2 : (1), a, bc, abc

(C) Block 1 : (1), ab, ac, bc
Block 2 : a, b, c, abc

(D) Block 1 : a, c, bc, ab
Block 2 : (1), ac, b, abc

78. In a k × k Latin Square design one observation is missing and suppose that R, C and T are totals of the known observations of row, column and treatment containing the missing value, and G is the total of $(k^2 - 1)$ available values. The estimated value of the missing observation is :

(A) $\dfrac{(k-1)(R+C+T)-G}{(k-1)(k-2)}$

(B) $\dfrac{k(R+C+T)-2G}{(k-1)(k-2)}$

(C) $\dfrac{(k-1)(R+C+T)-2G}{k^2-1}$

(D) $\dfrac{k(R+C+T)-G}{k^2-1}$

79. Consider a 2^3 partially confounded design involving 3 factors A, B and C with 4 replicates and 5 such repetitions. If the interactions AB, AC, BC and ABC are partially confounded each with 5 repetitions, the degrees of freedom for the error will be :

(A) 159

(B) 39

(C) 113

(D) 28

80. A strip-plot experiment is conducted with 5 levels of irrigation and 4 levels of an insecticide for foliar spray. The experiment contained 3 replications. The error degrees of freedom for the interaction effect will be :

(A) 24

(B) 19

(C) 22

(D) 30

81. The control chart suitable for the number of bubbles observed per square meter in a sheet of glass is :

(A) R-chart

(B) C-chart

(C) p-chart

(D) $\bar{X}$-chart

82. Some points beyond the control limits on the both sides in the $\bar{X}$-chart indicates that :

(A) The process level is gradually changing

(B) There is systematic differences within the samples

(C) The variability in the quality has been increased

(D) The level of the process has been shifted

83. From a data set consisting of defective items for 40 samples each of size 50, the average fraction defective is calculated as 0.10. The upper and lower control limits for the p-chart are respectively :

(A) 0.23, – 0.04

(B) 0.13, – 0.04

(C) 0.35, – 0.03

(D) 0.23, – 0.03

84. R-charts are generally used :

(A) To improve the quality of individual units of production

(B) To know basic variability of quality characteristics

(C) To help in reducing the average fraction defectives

(D) To show the random variation in quality of lots which are produced for inspection

85. The control chart used to monitor attributes is _____.

(A) Range chart

(B) Mean chart

(C) Chart for number of defects

(D) Chart for fraction defectives

86. Acceptance sampling is not used when _____.

(A) The test is destructive

(B) The cost of 100% inspection is quite high

(C) The supplier's process capability is very high

(D) Although the supplier process is satisfactory but a program is needed for continuous monitoring

87. The maximum value of the average outgoing quality overall possible values of the proportion defective is called :

(A) Acceptable quality level

(B) Average outgoing quality limit

(C) Average outgoing quality level

(D) Lot tolerance protection defective

88. Which of the following statement is not true regarding the OC curve ?

(A) The OC curve plots the probability of accepting the lot for a range of proportion of defective items

(B) In an ideal OC curve, the probability of acceptance of a lot having proportion of defectives less than or equal to LTPD value is equal to unity

(C) A sample size and an acceptance number are required to draw an OC curve

(D) As OC curve represents some risk factors, the protection afforded to the consumer by various sampling plans may be compared by examining their OC curves

89. For a single sampling plan with lot size 500, sample size 50, acceptance number 3 and probability of acceptance 0.90, compute A.T.I. :

(A) 95

(B) 455

(C) 175

(D) 505

90. In a double sampling plan, n_1 and n_2 are respectively sizes of the first and second samples. If P_1 and P_2 are respectively probabilities of decision on the basis of first and second samples, the expected sample size for taking a decision is :

(A) $ASN = n_1P_1 + (n_1 + n_2)P_2$

(B) $ASN = n_1P_1 + (n_1 + n_2)(1 - P_1)$

(C) $ASN = n_1P_1 + n_2(1 - P_2)$

(D) $ASN = n_1(1 - P_1) + (n_1 + n_2)P_1$

91. Which of the following cross validation techniques is better suited for the time series data ?

(A) k – fold cross validation

(B) Leave-one-out cross validation

(C) Stratified shuffle split cross validation

(D) Forward chaining cross validation

92. In the theory of time series, shortage of certain consumer goods before the annual budget of the country is due to :

(A) Seasonal variation

(B) Secular trend

(C) Irregular variation

(D) Cyclical variation

93. Simple average method for finding out seasonal indices is good when the time series :

(A) Is free from trend

(B) Has no cyclic effect

(C) Has no trend and cyclical variation

(D) Has all components

94. The moving averages in a time-series can be made free from the influences of :

(A) Seasonal and cyclic variations

(B) Seasonal and irregular variations

(C) Trend and cyclical variations

(D) Trend and random variations

95. The equation $U_t = \alpha\beta^{\gamma^t}$ represents a :

(A) Logistic curve

(B) Modified exponential curve

(C) Exponential curve

(D) Gompertz curve

96. Ratio to trend method for studying seasonal variation yields better results when :

(A) The periods are of long duration

(B) Seasonal variation is not a constant factor of trend value

(C) Cyclical variations are either present or if present not so significant

(D) Trend is linear

97. It is desired to determine a trend curve in a time series by a moving average covering consecutive sets of 5 points (U_{-2}, U_{-1}, U_0, U_1, U_2) and suppose we regard a quadratic polynomial as providing a satisfactory approximation. What would be the trend value ?

(A) $\frac{1}{35}(-3U_{-2} + 12U_{-1} + 17U_0 + 12U_1 - 3U_2)$

(B) $\frac{1}{21}(-2U_{-2} + 7U_{-1} + 11U_0 + 7U_1 - 2U_2)$

(C) $\frac{1}{21}(-3U_{-2} + 6U_{-1} + 15U_0 + 6U_1 - 3U_2)$

(D) $\frac{1}{15}(-2U_{-2} + 4U_{-1} + 11U_0 + 4U_1 - 2U_2)$

98. Examine critically the following statements and identify the wrong one :

(A) Harmonic analysis decomposes a periodic time series data into sum of sinusoidal functions

(B) Harmonic analysis for studying cyclical variation in the time series mainly utilizes Taylor series

(C) To avoid huge calculations, harmonic analysis may be carried out with some trial periods of oscillation guessed from the graph of the series

(D) Harmonic analysis is used to estimate phase and amplitude of the cyclical variation

99. As compared to the ratio to moving average method, which of the following statement is wrong for the link relatives method of studying seasonal variations ?

(A) The average of link relatives contains both trend and cyclical components

(B) Trend can be eliminated by applying corrections only if the growth is of constant amount or constant rate

(C) The actual calculations involved are more complicated

(D) Utilizes data more completely

100. Which of the following statement is not correct regarding the variate difference method ?

(A) The variance of the random component can be estimated by this method

(B) It can measure the random component when the trend of the time series is smooth

(C) It is not suitable if there are oscillations in the given series

(D) It can determine the degree of the polynomial with which the trend of the series can be represented

SPACE FOR ROUGH WORK

SPACE FOR ROUGH WORK

SEAL