ENTRANCE EXAMINATION FOR ADMISSION, MAY 2011.

Ph.D. (STATISTICS)

COURSE CODE : 149

Register Number :

Signature of the Invigilator
(with date)

COURSE CODE : 149

Time : 2 Hours | Max : 400 Marks

Instructions to Candidates :

1. Write your Register Number within the box provided on the top of this page and fill in the page 1 of the answer sheet using pen.
2. Do not write your name anywhere in this booklet or answer sheet. Violation of this entails disqualification.
3. Read each question carefully and shade the relevant answer (A) or (B) or (C) or (D) in the relevant box of the ANSWER SHEET using HB pencil.
4. Avoid blind guessing. A wrong answer will fetch you –1 mark and the correct answer will fetch 4 marks.
5. Do not write anything in the question paper. Use the white sheets attached at the end for rough works.
6. Do not open the question paper until the start signal is given.
7. Do not attempt to answer after stop signal is given. Any such attempt will disqualify your candidature.
8. On stop signal, keep the question paper and the answer sheet on your table and wait for the invigilator to collect them.
9. Use of Calculators, Tables, etc. are prohibited.

1. The size of a most powerful level α randomized test is less than α implies

(A) The test is not a level α test
(B) The power of the test is unity
(C) The power of the test is less than 1
(D) The test is not a size α test

2. If W_1 and W_2 are the most powerful critical regions of size α_1 and α_2 then $\alpha_1 < \alpha_2$ implies

(A) $W_1 \supset W_2$
(B) $W_1 \subset W_2$
(C) $W_1 \cap W_2 = \phi$
(D) No comparison of W_1 and W_2 is possible

3. If a statistical test is consistent then

(A) the test is unbiased
(B) the test is uniformly most powerful
(C) power of the test $\rightarrow$ 1 as sample size $\rightarrow \infty$
(D) the test is a similar test

4. For which one of the following test problem UMP test exist

(A) H : $\mu = \mu_0$ vs $K: \mu \neq \mu_0$ in the case of N $(\mu, 1)$
(B) H : $\theta = \theta_0$ vs $K: \theta \neq \theta_0$ in the case Cauchy $(\theta, 1)$
(C) H: $\theta = \theta_0$ vs $K: \theta \neq \theta_0$ in the case Poisson with parameter θ
(D) H: $\theta = \theta_0$ vs $K: \theta \neq \theta_0$ in the case U $(0, \theta)$

5. Which one of the following is a false statement?

(A) A minimal sufficient statistic need not be unique.
(B) A complete sufficient statistic is minimal.
(C) A minimal sufficient statistic is a function of all the sufficient statistics.
(D) A minimal sufficient statistic is complete

6. If T is an estimator of θ then the condition $E(T_n) \rightarrow \theta$ and $Var(T_n) \rightarrow 0$ as $n \rightarrow \infty$ is a sufficient condition for

(i) T weakly consistent for θ.
(ii) T mean square consistent for θ.

(A) (i) and (ii) are true
(B) only (ii) is true
(C) only (i) is true
(D) both (i) and (ii) are false

7. Which one of the following is a true statement if g is any continuous function?

(A) If T is consistent for θ then $g(T)$ is consistent for θ.

(B) If T is unbiased for θ then $g(T)$ is unbiased for θ.

(C) If T is sufficient for θ then $g(T)$ is sufficient for θ.

(D) All the above three options one false.

8. An estimator T of θ is UMVUE of a θ if T is unbiased for θ and

(A) $\text{Var}(T)$ = Cramer-Rao bound for the variance of unbiased estimators of θ.

(B) $\text{Var}(T) \leq$ variance of all other unbiased estimators of $\theta \ \forall\theta$.

(C) $\text{Var}(T)$ = Chapman-Robin bound for unbiased estimators of θ.

(D) $\text{Var}(T)$ = Bhattacharya bound for unbiased estimators of θ.

9. Which one of the following distribution does not belong to the one parameter exponential family?

(A) $N(0, \sigma^2)$ (B) $N(\mu, 1)$

(C) Exponential with mean $1/\theta$ (D) Cauchy $(\theta, 1)$

10. Which one of following is a false statement regarding the minimum chi-square estimate?

(A) It is consistent

(B) Asymptotically normally distributed

(C) Asymptotically UMVUE

(D) It is sufficient

11. If $\hat{\theta}$ is the MLE of θ, which one of the following is false?

(A) $\hat{\theta}$ is consistent for θ and asymptotically unbiased.

(B) If g is differentiable function of θ then MLE of $g(\theta)$ is $g(\hat{\theta})$.

(C) Asymptotically $\hat{\theta}$ is the UMVUE of θ.

(D) $\hat{\theta}$ is sufficient for θ.

12. The shortest expected width confidence interval for θ based on a random sample of n observations from U (0, θ) is

(A) Right sided (B) Left sided

(C) Two sided (D) depends on the value of α

13. Which one of the following is not true regarding the properties of the LSE $\hat{\theta}$ of θ in the linear model $Y = X\theta + e$, $e \sim N(0, \sigma^2 I_n)$, $\theta = (\theta_1, ..., \theta_k)'$?

(A) $\hat{\theta}_i$ is unbiased for θ_i $\forall i$

(B) $\hat{\theta}_i$ has the smallest variance among all linear unbiased estimators of θ_i.

(C) $\hat{\theta}$ has the smallest generalized variance among all linear unbiased estimators of θ.

(D) $\hat{\theta}$ is UMVUE of θ

14. Jordan Hahn decomposition theorem says that any distribution function can be decomposed as

(A) infinitely many distributions

(B) infinitely many discrete distributions

(C) a discrete and a continuous distribution

(D) finitely many discrete distributions

15. The distribution function $F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \le x < \frac{1}{2} \\ \left(x - \frac{1}{4}\right), & \frac{1}{2} \le x < 1 \\ 1, & x \ge 1 \end{cases}$ is

(A) continuous

(B) neither discrete nor continuous

(C) discrete

(D) not a distribution function

16. Let $\{X_n\}$ be a sequence of independent random variables such that $P[X_n = 1] = \frac{1}{n}$; $P[X_n = 0] = 1 - \frac{1}{n}, n = 1,2,...$ Which of the following statements is true?

(A) X_n converges to zero in rth mean

(B) X_n converges to zero almost surely

(C) Both (A) and (B)

(D) Neither (A) nor (B)

17. If the probability generating function of a random variable is $\phi(s)$, then its variance is

(A) $\phi''(0) - \phi'(0)$

(B) $\phi''(1) + \phi'(1) - [\phi'(1)]^2$

(C) $\phi''(1) - [\phi'(1)]^2$

(D) $\phi''(1) - \phi''(0)$

18. Let X be a positive integer valued random variable. Then its expected value is given by

(A) $\sum_{k=1}^{\infty} P[X > k]$ (B) $\sum_{k=1}^{\infty} P[X \geq k]$ (C) $\sum_{k=1}^{\infty} P[X \leq k]$ (D) $\sum_{k=1}^{\infty} P[X = k]$

19. Consider the following bivariate probability distribution.

		Y	
		-1	1
X	-1	1/2	0
	1	0	1/2

The characteristic function of (X,Y) is

(A) $\frac{1}{2}\left[e^{i(t_1+t_2)} + e^{-i(t_1+t_2)}\right]$

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

(C) $2e^{-i(t_1+t_2)}$

(D) $2e^{i(t_1+t_2)}$

20. Strong Law of Large Numbers is based on the concept of

(A) Weak convergence

(B) Almost sure convergence

(C) Convergence in probability

(D) Convergence in r^{th} mean

21. Cauchy – Schwartz inequality is a particular case of

(A) Jensen's inequality

(B) Minkowski inequality

(C) Holder's inequality

(D) C_r inequality

22. Let $\{A_n\}$ be independent events with $P\{A_n\} = \frac{1}{2}\ \forall n \geq 1$. Then $P\{A_n$ occuring infinitely often$\}$ is

(A) 0

(B) 0.5

(C) 1

(D) Cannot be determined

23. The density function of a random variable X is given by

$f(x) = \begin{cases} kx^2, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$. Then P[X>3/4 | X>1/2] is

(A) 3/5 (B) 37/56 (C) 25/50 (D) 1/3

24. The variance of a degenerate random variable is

(A) Infinity (B) Zero

(C) One (D) Undeterminable

25. The transition probability matrix of a Markov Chain is a stochastic matrix, because

(A) it is a square matrix (B) having non-negative elements

(C) its row sum is unity (D) satisfying all (A), (B) and (C)

26. Suppose X(t) represents the maximum temperature at a particular place in (0,t) then the stochastic process $\{X(t);\ t \geq 0\}$ is

(A) discrete time continuous state space

(B) continuous time discrete state space

(C) discrete time discrete state space

(D) continuous time continuous state space

27. If $\{X(t);\ t \in T\}$ is a stochastic process such that for $t_1 < t_2 < \ldots < t_n < t$ $P[a \leq X(t) \leq b / X(t_1)= x_1,\ X(t_2)= x_2, \ldots,\ X(t_n)= x_n] = P[a \leq X(t)\ \ b / X(t_n)= x_n]$, then the process is

(A) Poisson (B) Yule process

(C) Markov (D) Birth and Death process

28. Let $\{X_n;\ n \geq 0\}$ be a Markov Chain with three states 0, 1, 2 and with transition probability matrix $\begin{pmatrix} 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 3/4 & 1/4 \end{pmatrix}$ with initial distribution $P[X_0 = i] = 1/3$ i = 0, 1, 2.

Then

$P[X_2 = 2, X_1 = 1 / X_0 = 2]$ is equal to

(A) 3/15 (B) 3/16

(C) 3/17 (D) 3/19

29. The two step transition probability matrix of a Markov Chain with states 0,1,2 is given by $\begin{pmatrix} 5/8 & 5/16 & 1/16 \\ 5/16 & 1/2 & 3/16 \\ 3/16 & 9/16 & 1/4 \end{pmatrix}$ and the initial distribution $P[X_0 = 0] = 1/3$. Then

$P[X_2 = 1, X_0 = 0]$ is equal to

(A) 5/48 (B) 5/47 (C) 5/46 (D) 5/41

30. The period d(i) of a return state i is

(A) $d_i = Min\{m / P_{ii}^{(m)} > 0\}$

(B) $d_i = Max\{m / P_{ii}^{(m)} > 0\}$

(C) $d_i = GCD\{m / P_{ii}^{(m)} > 0\}$

(D) $d_i = LCM\{m / P_{ii}^{(m)} > 0\}$

31. Let F_{jk} denote the probability that starting with state j, the system will ever reach state k. A state j is said to be persistent if

(A) $F_{jj} = 1$ (B) $F_{jj} < 1$ (C) $F_{jj} > 1$ (D) $F_{jj} = 0$

32. The Chapman-Kolmogorov equation is useful for computing

(A) Periodicity of a Markov Chain

(B) Order of a Markov Chain

(C) Return state probability

(D) Higher order transition probability

33. Which one of the following is not the property of a Poisson process?

(A) Poisson Process is a Markov Process such that the conditional probabilities are constant

(B) It is a stationary process.

(C) The sum of two independent Poisson process is also a Poisson process.

(D) Poisson process has independent as well as stationary increments.

34. Suppose that customer arrives at a service counter in accordance with a Poisson process with mean rate of 2 per minute. The probability that the interval between two successive arrivals is more than 1 minute is

(A) e^{-1} (B) e^{-2} (C) e^{-3} (D) e^{-4}

35. If $\{N(t)\}$ is a Poisson process and s < t, then P[N(s) = k / N(t) = n] is

(A) Binomial distribution with p = s/t and q = 1 – s/t

(B) Binomial distribution with p = 1- s/t and q = s/t

(C) Poisson distribution with parameter λ t

(D) Poisson distribution with parameter λ s

36. If $\{N(t)\}$ is a Poisson process then the auto correlation between N(t) and N(t+s) is

(A) $\left(\frac{t+s}{t}\right)^{1/2}$ (B) $\left(\frac{t}{t+s}\right)^{1/2}$ (C) $\left(\frac{s}{t}\right)^{1/2}$ (D) $\left(1+\frac{t}{s}\right)^{1/2}$

37. If the probabilistic structure of the process is invariant under translation of the time axis then the process is called as

(A) Gaussian process (B) Poisson process

(C) Birth and Death process (D) Stationary process

38. Which one of the following is not the property of covariance function C(s,t) = cov[X(s),X(t)] of a stochastic process $\{X(t);\ t \geq 0\}$?

(A) It is symmetric

(B) It is non-negative definite

(C) The sum and product of two covariant functions are again a covariant function

(D) The reciprocal of covariance function is again a covariance function.

39. In a RBD the treatments are tested with the F ratio with df (4, 20). The total number of observations considered in this design is

(A) 24 (B) 80 (C) 30 (D) 32

40. Consider a completely randomized design with 4 treatments and 6 replications each. It is found that the total sum of square is 1300; treatment sum of square is 700. Then the mean square error is

(A) 20 (B) 30 (C) 40 (D) 50

41. In a Latin square design of order 5 with two missing values it is found that the error sum of square is 120. Then the mean square error is

(A) 10 (B) 12 (C) 15 (D) 8

42. A researcher conducts each replicate of a 2^7 Factorial experiment in 16 blocks of size 8 each. The number of interactions to be confounded is

(A) 8 (B) 7 (C) 15 (D) 16

43. For a Balanced Incomplete Block Design with parameters (v, b, r, k, λ), which of the design given below is not a BIBD?

(A) 6,10,5,3,2 (B) 7,7,3,3,1 (C) 8,14,7,4,2 (D) 7,7,4,4,2

44. A block design with v treatments in b blocks is said to be connected if the rank of the design matrix C is equal to

(A) b-1 (B) b (C) v-1 (D) v

45. A population consists of 10 students. The mark obtained by one student is 10 less than the average of the marks obtained by the remaining 9 students. Then the variance of the population of marks (σ^2) will always satisfy

(A) $\sigma^2 \geq 10$ (B) $\sigma^2 = 10$ (C) $\sigma^2 \leq 10$ (D) $\sigma^2 \geq 9$

46. With usual notation finite population correction is

(A) $(N-1)/n$ (B) $(N-n)/N$ (C) $(N-n)/n$ (D) $1-(1/n)$

47. In samples of moderate size the distribution of the ratio estimate shows a tendency to

(A) symmetry
(B) negative skewness
(C) non-normal
(D) positive skewness

48. Variance of the regression estimate is smaller than that of the mean per unit if

(A) $\rho = 0$ (B) $\rho \neq 0$ (C) $S_y^2 = 1$ (D) $S_x^2 = 1$

49. Systematic sampling means

(A) selection of n contiguous units
(B) selection of n units situated at equal distances
(C) selection of n largest units
(D) selection of n middle units in a sequence.

50. Let X be a r.v. denoting failure time of a component. Failure rate of the component is constant if and only if p.d.f. of X is

(A) Exponential
(B) Negative binomial
(C) Weibull
(D) Normal

51. The hazard rate of Weibull distribution is monotonic and it has increasing(decreasing) failure rate when the shape parameter (β) is

(A) $\beta>(<)1$ (B) $\beta<(>)1$ (C) $\beta>(=)1$ (D) $\beta=(<)1$

52. A two component system in parallel set up has constant failure rate. What is the reliability function for such configuration?

(A) R(t)= 1- [exp(-λ_1t) + exp(-λ_2t)]
(B) R(t)= exp(-λ_1t) + exp(-λ_1t) + exp(-(λ_1+ λ_2) t)
(C) R(t)= (1- exp(-λ_1t))(1-exp(-λ_2t))
(D) R(t)= 1+ (1- exp(-λ_1t))(1-exp(-λ_2t))

53. Sixty items were placed on test and the test was terminated after the 10 items failed. The failure time (in hours) were recorded as follows: 85, 151, 280, 376, 492, 520, 623, 715, 820 and 914. Assume that the failure time distribution is exponential with parameter λ. The estimated reliability at t = 600 is

(A) 0.94 (B) 0.88 (C) 0.75 (D) 0. 66

54. Under censoring scheme, a batch of transistors are kept for the life test in a random experiment and the test is terminated at a pre specific time t_c. This test experiment is called

(A) Type-I censoring (B) Type-II censoring

(C) Random censoring (D) Progressive censoring

55. For a maximization type linear programming problem, the simplex method is terminated when all values of

(A) $z_j - c_j \geq 0$ (B) $z_j - c_j \leq 0$ (C) $z_j - c_j = 0$ (D) $z_j - c_j \neq 0$

56. The feasible solution to a Linear Programming Problem

(A) must satisfy all of the problem's constraints simultaneously

(B) need not satisfy all of the constraints, only some of them

(C) must be a corner point of the feasible region

(D) must optimize the value of the objective function

57. If the constraints of the primal problem are all equations then the dual variable will be

(A) Positive (B) Negative

(C) Unrestricted in sign (D) Equal to zero

58. The initial solution of a transportation problem can be obtained by applying any known method. However, the only condition is that

(A) the solution is optimal (B) the rim conditions are satisfied

(C) the solution should not be degenerate (D) all the above

59. Dual Simplex method is used to solve a LPP when

(A) The solution is non-optimum and infeasible

(B) The solution is optimum and infeasible

(C) The solution is non-optimum and feasible

(D) The solution is optimum and degenerate

60. In post-optimality analysis of LPP, the feasibility condition is affected due to

(A) Changes in the cost vector

(B) Changes in the requirement vector

(C) Addition and deletion of linear constraints

(D) Both (B) and (C)

61. When there are more than one server, customer behaviour in which he moves from one queue to another is

(A) Balking (B) Jockeying (C) Reneging (D) Alternating

62. For a two-person zero-sum game with A and B and payoff matrix for player A, the optimum strategies are

(A) Minimax for A and Maximin for B

(B) Maximax for A and Minimax for B

(C) Minimin for A and Maximin for B

(D) Maximin for A and Minimax for B

63. The expected number of customers in the system in (M/M/1) : (∞/FCFS) queuing model is

(A) $\frac{\lambda}{\mu}$ (B) $\frac{\lambda}{\mu^2}$ (C) $\frac{\lambda}{(\mu - \lambda)}$ (D) $\frac{\mu}{(\lambda - \mu)}$

where λ and μ are mean arrival rate and mean service rate respectively.

64. If $X_1, X_2, ..., X_N$ is a random sample from $N_p(\mu, \Sigma)$ then the maximum likelihood estimators of mean vector μ and variance-covariance matrix Σ are

(A) $\overline{X}, \frac{A}{N}$ (B) $\overline{X}, \frac{A}{N-1}$ (C) $\overline{X}, \frac{S}{N}$ (D) $\overline{X}, \frac{S}{N-1}$

where A is Sum of squares and cross product matrix and S is sample covariance matrix.

65. If $X_1, X_2, ..., X_{N1}$ is a random sample from $N_p(\mu^{(1)}, \Sigma_1)$ and $X_1, X_2, ..., X_{N2}$ is a random sample from $N_p(\mu^{(2)}, \Sigma_2)$ then the test statistic for testing the hypothesis H: $\mu^{(1)} = \mu^{(2)}$ against K: $\mu^{(1)} \neq \mu^{(2)}$ where $\Sigma_1 = \Sigma_2 = \Sigma$ (unknown) is

(A) $T^2 = \frac{N_1 N_2}{N_1 + N_2 - 2}(\overline{X}^{(1)} - \overline{X}^{(2)})^T S^{-1}(\overline{X}^{(1)} - \overline{X}^{(2)})$

(B) $T^2 = N_1 N_2(\overline{X}^{(1)} - \overline{X}^{(2)})^T S^{-1}(\overline{X}^{(1)} - \overline{X}^{(2)})$

(C) $T^2 = \frac{N_1 N_2}{N_1 + N_2}(\overline{X}^{(1)} - \overline{X}^{(2)})^T S^{-1}(\overline{X}^{(1)} - \overline{X}^{(2)})$

(D) $T^2 = \frac{N_1 + N_2}{N_1 N_2}(\overline{X}^{(1)} - \overline{X}^{(2)})^T S^{-1}(\overline{X}^{(1)} - \overline{X}^{(2)})$

where S is the pooled sample variance-covariance matrix.

66. If $\Sigma = 1 = I$ and $p = 1$ in Wishart distribution then this distribution reduces to which univariate distribution?

(A) χ^2 distribution
(B) Student's t distribution
(C) Snedecor's F distribution
(D) Normal distribution

67. If the p-component vector Y is distributed according to N $(0, \Sigma)$ and Σ is non-singular then $Y\Sigma^{-1}Y$ is distributed according to

(A) χ^2 distribution with p degrees of freedom

(B) χ^2 distribution with (p-1) degrees of freedom

(C) $N_p(0, Y\Sigma Y)$

(D) $N_p(0, Y\Sigma^{-1}Y)$

68. The relation between Hotelling T^2 statistic and Mahalanobis D^2 statistic for testing equality of two mean vectors of normal population is

(A) $T^2 = \left(\frac{N_1 + N_2}{N_1 N_2}\right)D^2$
(B) $T^2 = \left(\frac{N_1 N_2}{N_1 + N_2}\right)D^2$

(C) $T^2 = \left(\frac{N_1 + N_2 - 2}{N_1 N_2}\right)D^2$
(D) $T^2 = \left(\frac{N_1 N_2}{N_1 + N_2 - 2}\right)D^2$

69. Consider the multiple regression model $Y = X\beta + \varepsilon$ where Y_{nx1}, $X_{nx(p+1)}$, $\beta_{(p+1)x1}$ and ε_{nx1}, $\varepsilon \sim N(0, \sigma^2 I)$. The m.l.e of β, $\hat{\beta} = (X'X)^{-1}X'Y$ is distributed according to

(A) $N(\beta, \sigma^2 I)$
(B) $N(\beta, I)$

(C) $N(\beta, \sigma^2(X'X)^{-1})$
(D) $N(\beta, \sigma^2(X'X))$

70. Characteristic function of Binomial distribution with parameters n and p is

(A) $(q+pe^{it})^n$
(B) $(p+qe^{it})^n$
(C) $(1+pe^{it})^n$
(D) $(p+e^{it})^n$

71. If the independent random variables X~ B $(3, \frac{1}{3})$ and Y~ B$(5, \frac{1}{3})$, then P[X+Y$\geq$1] is equal to

(A) $\left(\frac{2}{3}\right)^8$
(B) $1-\left(\frac{2}{3}\right)^8$
(C) $1+\left(\frac{2}{3}\right)^8$
(D) 1

72. If X is a poisson variate such that P(X=2)=9P(X=4)+90P(X=6) then the mean of X is

(A) 1
(B) 2
(C) 3
(D) 4

73. If two independent random variables X_1 and X_2 have the same geometric distribution, then X_1 given $(X_1 + X_2) = n$ follows

(A) Uniform distribution
(B) Poisson distribution
(C) Gamma distribution
(D) Hypergeometric distribution

74. If X has uniform distribution in [0,1], then pdf of Y=-2log X is given by

(A) $\frac{e^{-\frac{y}{2}}}{2}$
(B) $e^{-\frac{y}{2}}$
(C) $-\frac{y}{2}$
(D) $\frac{y}{2}$

75. The difference of two Poisson variates is

(A) Binomial variate
(B) Poisson variate
(C) Gamma variate
(D) None of the above

76. If X~ $\beta_2(\mu,\nu)$ then Y=$\frac{1}{1+X}$ follows

(A) $\beta_1(\mu,\nu)$
(B) $\beta_2(\mu,\nu)$
(C) $\beta_1(0,\nu)$
(D) $\beta_2(\mu,0)$

77. Ratio of two independent standard normal variates is

(A) Binomial variate
(B) Poisson variate
(C) Normal variate
(D) Standard Cauchy variate

78. If X~ χ^2_n, then X/2 follows Gamma distribution with parameter

(A) $\frac{n}{2}$
(B) $\frac{1}{2}$
(C) 1
(D) 2

79. Standard error of sample variance is

(A) $\sigma^2\sqrt{\frac{2}{n}}$
(B) $\sigma^2\sqrt{\frac{4}{n}}$
(C) $\sqrt{\frac{2}{n}}$
(D) σ^2

80. The cumulative distribution function of the smallest order statistic is

(A) $1-[1-F(x)]^n$
(B) $[1-F(x)]^n$
(C) 1-[1-F(x)]
(D) 1-F(x)

81. If X~ B(n,p), then $\text{Cov}\left(\frac{X}{n},\frac{n-X}{n}\right)$ is

(A) $\frac{-pq}{n}$
(B) $\frac{pq}{n}$
(C) $\frac{-p}{n}$
(D) $\frac{-q}{n}$

82. If X and Y have the same distribution, it does not imply that they are

(A) identical (B) different (C) independent (D) similar

83. Given that we have collected pairs of observations on two variables X and Y, we would consider fitting a straight line with X as an explanatory variable if

(A) the change in Y is an additive constant.

(B) the change in Y is a constant for each unit change in X

(C) the change in Y is a fixed percent of Y

(D) the change in Y is exponential

84. The least squares regression line is the line

(A) which is determined by use of a function of the distance between the observed Y 's and the regression line.

(B) which has the smallest sum of the squared residuals of any line through the data values.

(C) for which the sum of the residuals about the line is zero.

(D) which has all of the above properties

85. A botanist investigates the relationship between Y , the heights of seedlings (in inches), and X, the number of weeks after planting. The summary data are: n = 6, $\overline{X} = 4.67$, $\overline{Y} = 9.467$, $\sum X^2 = 154$, $\sum Y^2 = 696.54$, $\sum XY = 325.9$ The fitted regression line for seedling height on the number of weeks after planting is:

(A) $\overline{Y} = -2.8 + 2.62X$ (B) $\overline{Y} = 2.8 + 2.62X$

(C) $\overline{Y} = 2.62 + 2.8X$ (D) $\overline{Y} = 9.5 + 2.62X$

86. If you determine a point is an outlier, you would

(A) ignore it

(B) always delete it and then recalculate the line of best fit

(C) examine it carefully and try to determine what is causing it to be an outlier

(D) use it in the calculation of the Sum of Squared Errors

87. Neyman Pearson Lemma helps to obtain the most powerful test for testing

(A) a simple hypothesis against a simple alternative

(B) a simple hypothesis against a composite alternative

(C) a composite hypothesis against a composite alternative

(D) a composite hypothesis against a simple hypothesis.

88. If the points fit the regression line well

(A) the confidence interval of the slope of the regression line will be narrow

(B) the confidence interval of the intercept will be narrow

(C) the correlation coefficient will be numerically high

(D) all of these are true.

89. Compared with parametric equivalents, nonparametric tests are

(A) Less powerful

(B) Less likely to reject the null hypothesis

(C) More conservative

(D) All of these

90. What could have undue influence on the value of the correlation and regression coefficients estimated in fitting a model to the data set?

(A) Normality (B) Outliers (C) Regressors (D) Error term

91. For what type of data, we use Mc Nemar's test

(A) Ordinal (B) Interval (C) Ratio (D) Nominal

92. In Kruskal – Wallis ANOVA, if significance is observed then what is the suitable post-hoc test to be used

(A) Duncan's multiple range test

(B) Adjusted significance level with Bonferonni test

(C) Least significance test

(D) Scheffe's test

93. The connection between almost sure (a.s) convergence, convergence in probability (p) and in mean (m) is

(A) a.s $\Rightarrow$ m $\Rightarrow$ p (B) a.s $\Rightarrow$ p; p $\Rightarrow$ m

(C) a.s $\Rightarrow$ p; m $\Rightarrow$ p (D) m $\Rightarrow$ a.s $\Rightarrow$ p

94. X_1, X_2, X_3 are independent observations from a normal population with mean θ and variance unity.

Two statistics T_1 and T_2 are defined as $T_1 = \frac{X_1 + X_2 + 2X_3}{4}$ and $T_2 = \frac{X_1 + X_2 + X_3}{3}$. The efficiency of T_1 relative to T_2 is given by

(A) $\frac{8}{9}$ (B) $\frac{3}{4}$ (C) $\frac{9}{8}$ (D) $\frac{4}{3}$

95. The slope of the regression line

(A) may be positive or negative

(B) is the same as the correlation coefficient

(C) is indeterminate

(D) all of these

96. For a degenerate random variable at $a \in R$

(A) $P(X = a) = 0$ (B) $P(X = a) = \infty$ (C) $P(X = a) = -\infty$ (D) $P(X = a) = 1$

97. **Hypergeometric distribution tends to Binomial distribution as $N \to \infty$ and**

(A) $\frac{M}{N} \to q$ (B) $\frac{M}{N} \to p$ (C) $\frac{M}{N} \to 0$ (D) $\frac{M}{N} \to 1$

98. Variance is always less than mean in the

(A) Poisson distribution (B) Negative Binomial distribution

(C) Binomial distribution (D) Chi-square distribution

99. If X^2 and Y^2 are independent then

(A) X and Y are independent (B) X and Y need not be independent

(C) $E(X^2) = E(Y^2)$ (D) $E(XY) = 0$

100. In a stratified random sampling scheme the population characteristics can be more efficiently estimated from a stratified sample than from overall simple random sample if

(A) Strata means differ widely, and within strata variation is high.

(B) Strata means do not differ widely, and within strata variation is high.

(C) Strata means differ widely, and within strata variation is low.

(D) Strata means do not differ widely, and within strata variation is low.