STATISTICS

1. Let A and B be subsets of Ω. Then $\left[A \cap (\Omega - B)\right] \cup B$ is

(A) $A \cup B$
(B) $A \cap B$
(C) A
(D) B

2. If $f(x) = \begin{cases} 1, & \text{if } x \text{ is a rational number} \\ 0, & \text{if } x \text{ is an irrational number} \end{cases}$, then $f \circ f\left(\sqrt{4}\right)$ is equal to

(A) 0
(B) 1
(C) both 0 and 1
(D) 2

3. The intersection of finite collection of open sets is

(A) open
(B) closed
(C) empty
(D) unbounded

4. The set of real numbers is

(A) uncountable
(B) countable
(C) finite
(D) bounded

5. For set $A = \left\{ \frac{(-1)^n}{n^2}, n = 1, 2, \ldots \right\}$, inf A and sup A are equal to

(A) –1 and ¼
(B) –1 and ½
(C) 1 and ¼
(D) –2 and ½

Only for reference. Do not copy

6. Let $F_n = \left[-\frac{1}{n}, \frac{1}{n}\right]$, then $\bigcap_{n=1}^{\infty} F_n$ is a

(A) open
(B) closed
(C) empty
(D) unbounded

7. If $f(x) = \begin{cases} \sin\left(\frac{\pi x}{2}\right), x < 1 \\ |2x-3|, x \geq 1 \end{cases}$ then $f(x)$ is

(A) not continuous at $x = 1$
(B) not differentiable at $x = 1$
(C) continuous at $x = 1$
(D) differentiable at $x = 1$

8. The series $\sum_{n=1}^{\infty} \frac{1}{n(\log n)^p}$ is

(A) convergent if $p > 1$
(B) divergent if $p > 1$
(C) convergent if $p \leq 1$
(D) divergent if $p = 1$

9. For the sequence $\left\{\frac{2^n}{n}\right\}$,

(A) limit exists and is equal to 0
(B) limit does not exists
(C) limit exists and is equal to 1
(D) limit exists and is equal to 2

10. Eigen values of a square matrix are roots of

(A) differential equation
(B) logical equation
(C) characteristic equation
(D) integral equation

Only for reference. Do not copy

11. $\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$ is

(A) e^x
(B) ∞
(C) 1
(D) e^{-x}

12. The series $1+\frac{1}{2^2}+\frac{2^2}{3^3}+\frac{3^3}{4^4}+\ldots$ is

(A) convergent and equal to 1
(B) divergent
(C) convergent and equal to $1/2$
(D) convergent and equal to 0

13. If a series $\sum_{n=1}^{\infty} a_n$ converges, then a_n

(A) converges to 1
(B) diverges
(C) converges to any number
(D) converges to 0

14. Let $f(x)=\begin{cases}x^2+3x+a, x\le 1\\ bx+2, x>1\end{cases}$ be differentiable at $x=1$, then

(A) $a=1, b=4$
(B) $a=2, b=2$
(C) $a=5, b=3$
(D) $a=3, b=5$

15. Let $A=\begin{bmatrix}2 & 0 & 5\\ 1 & 2 & 3\\ -1 & 5 & 1\end{bmatrix}$. Then system of linear equations $AX= Y$ has a solution

(A) only for $Y=(x \quad 0 \quad 0)', x\in R$
(B) only for $Y=(0 \quad y \quad 0)', y\in R$
(C) for all $Y\in R^3$
(D) only for $Y=(0 \quad y \quad z)', y, z\in R$

Only for reference. Do not copy

16. Subject to conditions $0 \le x \le 1, 0 \le y \le 5,$ minimum value of function $f(x, y) = 4x - 5y + 10$ is

(A) 10
(B) 0
(C) –15
(D) –25

17. Let n be a fixed natural number. Then the series $\sum_{m \ge n} \frac{(-1)^m}{m}$ is

(A) absolutely convergent
(B) convergent
(C) absolutely convergent if $n > 100$
(D) divergent

18. Sum of series $\left(\frac{1}{2.3} + \frac{1}{2^2.3}\right) + \left(\frac{1}{2^2.3^2} + \frac{1}{2^3.3^2}\right) + \ldots + \left(\frac{1}{2^n.3^n} + \frac{1}{2^{n+1}.3^n}\right) \ldots$ is

(A) 3/8
(B) 3/10
(C) 3/14
(D) 3/16

19. A set of linear equations is represented by the matrix equation $Ax = b$. The necessary condition for the existence of a solution for this system is

(A) A must be invertible
(B) b must be linearly dependent on the columns of A
(C) b must be linearly independent on the columns of A
(D) det $(A) = 0$

20. If $|A|$ denotes the determinant of the matrix A then which of the following is not correct?

(A) $|A + B| = |A| + |B|$
(B) $|AB^{-1}| = \frac{|A|}{|B|}$
(C) $|AB| = |A||B|$
(D) $|A^{-1}| = \frac{1}{|A|}$

Only for reference. Do not copy

21. If A is an idempotent matrix, then

(A) Rank (A) < Rank (A^{10})
(B) Rank (A) > Rank (A^{10})
(C) Rank (A^2) = 2 Rank (A)
(D) Rank (A) = Rank (A^{10})

22. The function $f(x) = |x|$ is

(A) both continuous and differentiable at $x = 0$
(B) continuous at $x = 0$ but not differentiable at $x = 0$
(C) differentiable at $x = 0$ but not continuous at $x = 0$
(D) None of the above

23. The function $x^3 - 6x^2 + 9x + 25$ has

(A) a maxima at $x = 1$ and a minima at $x = 3$
(B) a maxima at $x = 3$ and a minima at $x = 1$
(C) no maxima but a minima at $x = 3$
(D) no minima but a maxima at $x = 1$

24. If A is a 4×4 matrix with $|A| = 3$ then $|3A|$ is

(A) 3
(B) 243
(C) 27
(D) 81

25. If $\log_4 8 = x$, then value of x is

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

Only for reference. Do not copy

26. Rank of an identity matrix of order n is

(A) $(n - 1)$
(B) n
(C) $(n + 1)$
(D) 1

27. System of equations $4x + 6y = 5$ and $8x + 12y = 10$ has

(A) a unique solution
(B) no solution
(C) infinitely many solutions
(D) two solutions

28. Value of $\lim_{x \to 0} \frac{5x^4 + 8x^2}{3x^4 - 16x^2}$ is

(A) $3/2$
(B) $-1/2$
(C) 1
(D) 2

29. The sequence $\left\{ \frac{\sin n\frac{\pi}{2}}{n} \right\}_{n=1}^{\infty}$ is :

(A) divergent
(B) convergent
(C) neither convergent nor divergent
(D) none of the above

30. Let $X \sim U(0, 1)$. If

(i) $P(X \geq 0.27) = 0.73$
(ii) $P(0.27 \leq X \leq 1.27) = 0.73,$ then

(A) (i) is true
(B) (ii) is true
(C) both (i) and (ii) are true
(D) neither (i) nor (ii) is true.

Only for reference. Do not copy

31. The quadratic form $Q(x, y, z) = -12xy - 6xz + 10x^2 + 4z^2$ is

(A) positive definite
(B) positive semidefinite
(C) negative definite
(D) indefinite

32. The function $f(x) = \begin{cases} x \text{ for } x \le 1 \\ 3x^2 \text{ for } x > 1 \end{cases}$ is

(A) continuous at $x = 1$
(B) not continuous at $x = 1$
(C) differentiable at $x = 1$
(D) neither continuous nor differentiable at $x = 1$

33. Let A and B be two events such that $P(A) = 1/4, P(B) = 1/3$ and $P(A \cap B) = 1/12$. If

(i) $P(A \cap \bar{B}) = 1/4$
(ii) $P(\bar{A} \cap \bar{B}) = 1/2$
(iii) $P(B \mid \bar{A}) = 1/3$

(A) both (ii) and (iii) are true
(B) both (i) and (ii) are true
(C) both (i) and (iii) are true
(D) All are true

34. A person goes to work by using three routes. The probabilities that he arrives on time using routes I, II and III are 50%, 65% and 71% respectively. If he is equally likely to choose any of the routes, and arrives at work on time, then the probability of choosing route I is

(A) 32.3%
(B) 26.9%
(C) 45.7%
(D) 56.2%

35. Which of the following distributions can be considered as a generalization of geometric distribution?

(A) Poisson
(B) Hyper-geometric
(C) Binomial
(D) Negative binomial

Only for reference. Do not copy

36. The probability of Radha living 20 years more is $\frac{1}{5}$ and that of Lakshmy is $\frac{1}{7}$. Then the probability that atleast one of them will survive 20 years hence is:

(A) $\frac{1}{35}$
(B) $\frac{11}{35}$
(C) $\frac{12}{35}$
(D) $\frac{13}{35}$

37. The distribution which has mean always less than variance is:

(A) Beta distribution
(B) Lognormal distribution
(C) Negative Binomial distribution
(D) Weibull distribution

38. Which of the following is TRUE?

(A) p-value measures the probability that the null hypothesis is true
(B) Extremely small p-value indicates that actual data differ significantly from expected data when null hypothesis is true
(C) Large p-value is an evidence against the null hypothesis
(D) p-value measures the probability of Type II error

39. The relation between the mean and variance of chi-square distribution with '*n*' degrees of freedom is :

(A) variance = mean / 2
(B) variance = 2.mean
(C) variance = mean
(D) variance = $(\text{mean})^2$

Only for reference. Do not copy

40. Suppose a sample of size $n = 100$ is collected from some population and the data are used to calculate a 95% confidence interval for the population mean. Now suppose the sample size is increased to $n = 200$. Keeping all else constant, which of the following is expected to occur as a result of increase in the sample size?

I The standard error would decrease.
II Width of the 95% confidence interval would increase.
III The margin of error would decrease.
IV The margin of error would increase.

(A) II and III
(B) I and III
(C) I and II
(D) I, II, and IV

41. Approximately what percentage of scores fall within two standard deviations of the mean in a normal distribution?

(A) 34%
(B) 99%
(C) 95%
(D) 68%

42. If a test was generally very easy, except for a few students who had very low scores, then the distribution of scores would be

(A) positively skewed
(B) negatively skewed
(C) not skewed at all
(D) normal

43. If the mean and variance of a uniformly distributed random variable are 3 and 1/3 then the distribution is defined over

(A) [2, 3]
(B) [3, 4]
(C) [0, 1]
(D) [2, 4]

44. Let $P(A) = .43, P(B) = .4$ and $P(A \cap B) = .76$. Then A and B are

(A) disjoint events
(B) equally likely events
(C) independent events
(D) dependent events

Only for reference. Do not copy

45. If the Rank of the matrix $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & x & 0 \\ 6 & 7 & 8 & 9 \end{bmatrix}$ is 3 then the value of x is

(A) 6
(B) 1
(C) –6
(D) 0

46. A method currently used by doctors to screen women for possible breast cancer fails to detect cancer in 15% of the women who actually have the disease. A new method has been developed that researchers hope will be able to detect cancer more accurately. A random sample of 80 women known to have breast cancer is to be screened using the new method. At the 0.05 level of significance, the researchers will be able to conclude that the new method is better than the one currently in use if the appropriate test statistic has a value:

(A) greater than 1.96
(B) less than 1.645
(C) less than –1.645
(D) greater than –1.96

47. An unbiased estimator of population mean is

(A) sum of all sample observations
(B) sample median
(C) mean of sample observations
(D) sample standard deviation

48. Given $r_{12} = 0.6$, $r_{13} = 0.5$ and $r_{23} = 0.8$, the value of $r_{12.3}$ is

(A) 0.40
(B) 0.72
(C) 0.38
(D) 0.47

49. A medical researcher wishes to do a random survey of 100 female doctors and 100 male doctors. Which sampling technique should he use?

(A) Stratified
(B) Systematic
(C) Simple random
(D) Cluster

Only for reference. Do not copy

50. In a random sample of size 3 drawn from $N(0,\sigma^2)$, which of the following estimators is not unbiased for σ^2 ?

(A) X_1^2

(B) $(X_1 - X_2)^2$

(C) $\dfrac{(X_1 - X_2)^2}{2}$

(D) $\dfrac{2X_1^2 - X_2^2 + 5X_3^2}{6}$

51. Given the two lines of the regression as, $3X - 4Y + 8 = 0$ and $4X - 3Y = 1$, the means of X and Y are:

(A) $\bar{X} = 4, \bar{Y} = 5$

(B) $\bar{X} = 2, \bar{Y} = 5$

(C) $\bar{X} = 4, \bar{Y} = 2$

(D) $\bar{X} = 3, \bar{Y} = 4$

52. If the correlation coefficient between x and y is zero, then the angle between two lines of regression is

(A) 70°
(B) 90°
(C) 60°
(D) 30°

53. If a constant 20 is subtracted from the values of X and Y, the regression coefficient of Y on X is

(A) 1/20th of the original regression coefficient
(B) reduced by 20
(C) not changed
(D) increased by 20

Only for reference. Do not copy

54. The level of significance is the probability of

(A) rejecting true null hypothesis
(B) failing to reject true null hypothesis
(C) rejecting false null hypothesis
(D) failing to reject false null hypothesis

55. Analysis of variance utilises

(A) t-test
(B) F-test
(C) Chi-Square test
(D) Z-test

56. If the unit of measurement of a commodity changes, the value of index number

(A) also changes
(B) remains same
(C) increases
(D) decreases

57. POSTHOC tests are used for

(A) pairwise comparison of means
(B) equality of three population means
(C) equality of two population variances
(D) equality of three or more population means

58. The hypothesis that the population variance has a specified value is tested by using

(A) F-Test
(B) Z-Test
(C) Chi-Square Test
(D) t-Test

59. Gender of babies born in a hospital falls under

(A) ratio scale
(B) ordinal scale
(C) interval scale
(D) nominal scale

Only for reference. Do not copy

60. All the diagonal elements of a skew symmetric matrix are

(A) 1
(B) 0
(C) 2
(D) any integer

61. For a positively skewed distribution,

(A) mean is on the right of mode
(B) mean is on the left of mode
(C) mean is same as mode
(D) mean is same as median

62. If level of significance is reduced, then the size of the rejection region is

(A) reduced
(B) increased
(C) unaltered
(D) depends upon the form of alternative hypothesis

63. For the data set 4, 3, 2, 5, 4, 2, 2, 5, 6, 7, 1, 3, 8, if one of the values with magnitude 5 is changed to 6, then which measure of central tendency will change?

(A) Mean
(B) Mean and median
(C) Mean and mode
(D) Median

64. In Non-parametric theory, which measure of location is most frequently used?

(A) Mean
(B) Mode
(C) Median
(D) Harmonic Mean

65. The t-distribution is similar to the standard normal distribution in all but one of the following characteristics. Which one is it?

(A) Symmetrical
(B) Bell-shaped
(C) Continuous
(D) It has mean = 0 and variance = 1

Only for reference. Do not copy

66. Which of the following are used for visual examination of the relationship between two quantitative variables?

(A) Bar graphs
(B) Pie graphs
(C) Line graphs
(D) Scatterplots

67. Which non-parametric test is used to test the equality of mean of more than two samples when the assumptions for parametric tests fail?

(A) Run Test
(B) Kruskal Wallis Test
(C) Mann Whitney Test
(D) Wilcoxon Signed Rank Test

68. Which test is used to test the normality of the data?

(A) Run Test
(B) Kruskal Wallis Test
(C) Mann Whitney Test
(D) Kolmogorov Smirnov Test

69. If β is the regression coefficient, then to test $H_0 : \beta = \beta_0$, against the alternative $H_1 : \beta \neq \beta_0$, the test statistics used under H_0 has

(A) t distribution
(B) F distribution
(C) Chi-square distribution
(D) Normal distribution

70. For testing $H : \theta = 1$ against $K : \theta = 2$ based on a single observation drawn from Poisson(θ) the test "Reject H if $x > 1$" is to be used. The probability of type-I error is

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

Only for reference. Do not copy

71. Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}$, then

(A) all eigen values are equal
(B) all eigen values are distinct
(C) two eigen values are same but third one is different
(D) eigen values of A are –1, 0 and +1

72. For a skewed distribution with median = 16.2 and mode = 20.4, mean is approximately

(A) 12
(B) 14.1
(C) 13.6
(D) 11.4

73. If the correlation coefficient between the variables x and y is negative, then the regression coefficient of y on x is

(A) negative
(B) positive
(C) zero
(D) greater than one

74. Let $\sin xy = x$, then $\frac{dy}{dx}$ is

(A) $\frac{1}{y}\left(\frac{1}{\cos xy} - x\right)$
(B) $\frac{1}{x}\left(\frac{1}{\cos xy} - y\right)$
(C) $\left(\frac{1}{\cos xy} - x\right)$
(D) $\left(\frac{1}{\cos xy} - y\right)$

Only for reference. Do not copy

75. Let $X_1, X_2, ..., X_n$ be independent random variables each $N(\theta, 1)$ where θ is unknown. Then maximum likelihood estimator of θ is

(A) $\bar{X}$

(B) $\sum_{i=1}^{n} X_i$

(C) $\prod_{i=1}^{n} X_i$

(D) $\sum_{i=1}^{n} X_i^2$

76. Suppose in a certain population, 20 percent are immune to some disease. If a random sample of size 5 is selected from the population, then the probability that exactly two are immune in the sample is

(A) 0.2048
(B) 0.3124
(C) 0.5421
(D) 0.2134

77. If X is distributed as 't' with 10 degrees of freedom, then which of the following statements is not correct?

(A) $E(X^{11}) = 0$

(B) $E(X^9) = 0$

(C) $\text{var}(x) = \frac{10}{7}$

(D) Mean and variance are not equal

78. When sampling is from a Normally distributed population, then which of the following is false about the distribution of the sample mean?

(A) It is Normal
(B) It has same mean as the population
(C) It has variance equal to the population variance divided by the sample size
(D) It has same variance as the population

Only for reference. Do not copy

79. If population size is 10 and samples of size 4 are drawn without replacement, then the number of possible samples is

(A) 200
(B) 250
(C) 210
(D) 220

80. Which of the following is NOT true for t-distribution?

(A) It is symmetrical about mean
(B) It always has variance greater than one
(C) It is less peaked than Normal distribution
(D) t-distribution approaches Normal distribution for large sample size

81. $\int_1^e \left(\frac{x^2 - 1}{x} \right) dx$ is equal to

(A) $\frac{e^2}{2}$

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

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

(D) $\frac{e^2}{2} - \frac{3}{2}$

82. If V_{ran}, V_{prop} and V_{opt} respectively denote the variance of the estimator of the population mean under SRSWOR, stratified sampling under proportional and Neymann allocation, then

(A) $V_{ran} > V_{prop} > V_{opt}$
(B) $V_{ran} < V_{prop} < V_{opt}$
(C) $V_{ran} < V_{prop}$ and $V_{prop} < V_{opt}$
(D) $V_{ran} < V_{prop}$ and $V_{prop} > V_{opt}$

Only for reference. Do not copy

83. The function $f(x) = x^x$ attains minimum at

(A) e

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

(C) $\frac{-1}{e}$

(D) $-e$

Only for reference. Do not copy

84. If A, B and C are three independent events, then probability of their joint occurrence is

(A) $P(A)+P(B)+P(C)$
(B) $P(A).P(B).P(C)$
(C) $P(A)-P(B)-P(C)$
(D) $1/\left[P(A).P(B).P(C)\right]$

85. If A and B are square matrices of order $n \times n$, then which of the following is not true?

(A) $\det(AB)=\det(A).\det(B)$
(B) $\det(kA)=k^n \det(A)$
(C) $\det(A+B)=\det(A)+\det(B)$
(D) $\det(A^T)=1/\det\left(A^{-1}\right)$

86. Which sampling design is most appropriate for cluster sampling?

(A) Simple random sampling without replacement
(B) Simple random sampling with replacement
(C) Stratified random sampling
(D) Quota sampling

87. Which of the following statements is NOT true?

(A) In a symmetric distribution, the mean and the median are equal
(B) The first quartile is equal to the twenty-fifth percentile
(C) In a symmetric distribution, the median is halfway between the first and the third quartiles
(D) The median is always greater than the mean

88. A fair coin is tossed repeatedly. If tail appears on first four tosses of a coin, the probability that head will appear on fifth toss is

(A) 1/2
(B) 1/32
(C) 3/32
(D) 1/4

Only for reference. Do not copy

89. In tossing three coins, the probability of at most one head is

(A) 1/2
(B) 7/8
(C) 3/8
(D) 3/4

90. In a city, 50% read newspaper I, 35% read newspaper II, 30% read newspaper III, 20% read I and II, 25% read I and III and 10% read II and III. I, II and III are read by 14%. The percentage of people who do not read any of the newspapers is

(A) 74
(B) 26
(C) 70
(D) 84

91. The type of sampling in which each member of the population selected for the sample is returned to the population before the next member is selected is called

(A) Sampling without replacement
(B) Sampling with replacement
(C) Simple random sampling
(D) Systematic sampling

92. Let $X_1, X_2, ..., X_6$ be independent random variables such that $P(X_i = -1) = P(X_i = 1) = \frac{1}{2}$, then $P\left(\sum_{i=1}^{6} X_i = 4\right)$ is

(A) 3/32
(B) 3/16
(C) 1/16
(D) 1/32

93. Suppose scores on a training post test are being predicted from number of years of education and the score on an aptitude test given before training. The regression equation is $Y = 25 + .5X_1 + 10X_2$, where X_1 = years of education and X_2 = aptitude test score. The predicted score for someone with 10 years of education and aptitude test score of 5 is

(A) 25
(B) 50
(C) 35
(D) 80

Only for reference. Do not copy

94. Suppose that a study is conducted on three different complications associated with a disease. The probabilities associated with three complications are given in the following table:

Complication 1	**Complication** 2	**Complication** 3
0.1	0.3	0.6

Suppose that a group of six patients is chosen randomly. Then probability of selecting 2, 1 and 3 patients with complication type 1, 2 and 3 respectively is

(A) 0.0365
(B) 0.0389
(C) 0.0256
(D) 0.0154

95. $\int_0^2 [x] dx^2$ is equal to

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

96. Let $f(x,y) = \begin{cases} \frac{(x+y)}{3}, 0 < x < 1, 0 < y < 2 \\ 0, \quad \text{otherwise} \end{cases}$. Then $E\left[3X(X+2Y)\right]$ is equal to

(A) $31/6$
(B) $9/2$
(C) $11/2$
(D) $29/6$

97. Cans of soft drinks cost \$0.30 in a certain vending machine. What is the expected value and variance of daily revenue (Y) from the machine, if X, the number of cans sold per day has $E(X) = 125$, and $\text{var}(X) = 50$?

(A) $E(Y) = 37.5,\ \text{var}(Y) = 50$
(B) $E(Y) = 37.5,\ \text{var}(Y) = 4.5$
(C) $E(Y) = 37.5,\ \text{var}(Y) = 15$
(D) $E(Y) = 37.5,\ \text{var}(Y) = 45$

Only for reference. Do not copy

98. If $x \geq 1$ is the critical region for testing $H_0 : \theta = 2$ against $H_1 : \theta = 1$. Based on a single observation from $f(x) = \theta e^{-\theta x}, x \geq 0, \theta > 0$, the probability of Type II error is

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

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

(C) $\frac{e-1}{e}$

(D) $\frac{1}{e^3}$

99. What would happen (other things equal) to a confidence interval if you calculated a 99 percent confidence interval rather than a 95 percent confidence interval?

(A) It will be narrower
(B) It will not change
(C) The sample size will increase
(D) It will become wider

100. In Completely Randomized Design, the value of F statistic is calculated by using

(A) $\frac{MS_{Total}}{MS_{Error}}$

(B) $\frac{MS_{Error}}{MS_{Total}}$

(C) $\frac{MS_{Treatments}}{MS_{Error}}$

(D) $\frac{MS_{Error}}{MS_{Treatments}}$

101. What is the standard deviation of a sampling distribution called?

(A) Sampling error
(B) Sample error
(C) Standard error
(D) Simple error

Only for reference, Do not copy

102. Let X be a continuous random variable having pdf $f(x) = \begin{cases} \frac{2}{x^3}, \text{for } x \geq 1 \\ 0, \quad \text{for } x < 1 \end{cases}$. Then

(A) expectation does not exist but variance exists
(B) both expectation and variance exist
(C) expectation exists but variance does not exist
(D) neither expectation nor variance exists

103. Let X, Y and $Z = (X + 2Y)$ be random variables such that $V(X) = 4$, $V(Y) = 3$ and $V(Z) = 4$. Then, which of the following statements are true?

(i) $\text{cov}(X, Y) = -3$

(ii) Correlation $(X, Y) = \frac{-\sqrt{3}}{2}$

(iii) Regression Coefficient of Y on $X = \frac{-3}{4}$

(iv) $\text{cov}(X, Z) = -2$

(A) Only (i) and (ii) are true
(B) All the statements are true
(C) Only (ii) and (iii) are true
(D) Only (i) and (iii) are true

104. Let X_1, X_2, X_3 and X_4 be uncorrelated random variables each with variance σ^2. Then correlation between $(X_1 + X_2 + X_3)$ and $(X_1 + X_2 + X_4)$ is

(A) 0
(B) 2/3
(C) 1/3
(D) 1

105. A statistical procedure used to compare 2 or more group means is known as

(A) One-way analysis of variance
(B) Post hoc test
(C) t-test
(D) Simple regression

Only for reference. Do not copy

106. The following layout

$A \quad B \quad C \quad D$
$A \quad C \quad B \quad D$
$B \quad A \quad C \quad C$
$A \quad A \quad B \quad C$

meets the requirements of a:

(A) completely randomized design
(B) randomized block design
(C) Latin square design
(D) CRD or RBD

107. The number of accidents per week in a small city has a Poisson distribution with mean equal to 3. The probability of exactly 2 accidents in 2 weeks is

(A) $2e^{-3}$
(B) $9e^{-3}$
(C) $18e^{-6}$
(D) $18e^{-3}$

108. Three numbers are chosen at random without replacement from the set $A = \{x \mid 1 \le x \le 10, x \in N\}$. The probability that the minimum of the chosen numbers is 4 and maximum is 7, is

(A) $\frac{1}{12}$
(B) $\frac{1}{15}$
(C) $\frac{1}{60}$
(D) $\frac{1}{40}$

109. The joint probability of events *A* and *B* is 32 percent with the probability of event *A* being 60 percent and the probability of event *B* being 50 percent. Based on this information, the conditional probability of event *A* given event *B* has occurred is closest to:

(A) 30.0%
(B) 53.3%
(C) 64.0%
(D) 46%

Only for reference. Do not copy

110. Suppose 65% of a herd of cattle is infected with a particular disease. Let X = the number of non-infected cattle in a sample of size 6. The distribution of X is

(A) binomial with $n = 6$ and $p = 0.65$
(B) binomial with $n = 6$ and $p = 0.35$
(C) the same as the distribution of X, the number of infected cattle
(D) Poisson with $\lambda = .65$

111. The distribution of phone calls arriving in one minute periods at a switchboard is assumed to be Poisson with parameter λ. During 100 periods, the following distribution was obtained:

No. of calls	0	1	2	3	4 or more
Frequency	30	43	21	6	0

An estimate for λ based on this data set is:

(A) 1.00
(B) 1.03
(C) 1.04
(D) 1.33

112. In simple random sampling without replacement from a population of size 100, samples of size 8 are to be selected. If $s^2 = .789$, then estimate of the variance of population mean is

(A) 0.653
(B) 0.891
(C) 0.0907
(D) 0.4938

113. In general, which of the following statements is FALSE?

(A) The sample mean is more sensitive to extreme values than the median
(B) The sample range is more sensitive to extreme values than the standard deviation
(C) The sample standard deviation is a measure of spread around the sample mean
(D) The coefficient of variation is the ratio of sample mean to the sample standard deviation

Only for reference. Do not copy

114. Let X be the number of phone calls received at a call center during business hours. And Y be the number of phone calls received at a call center outside business hours. Let X and Y be independent random variables following $P(\lambda_1)$ and $P(\lambda_2)$, $\lambda_1 \neq \lambda_2$ respectively. Then conditional distribution of X given $(X + Y)$ is

(A) Poisson
(B) Discrete Uniform
(C) Binomial
(D) Negative Binomial

115. Sum of the eigen roots of the matrix $A = \begin{bmatrix} a & 2 \\ 1 & b \end{bmatrix}$

(A) $a+b$
(B) ab
(C) $2a+b$
(D) $a+2b$

116. Forty students wrote a Statistics examination having a maximum of 50 marks. The mark distribution is given in the following stem-and-leaf plot:

0|28
1|2245
2|01333358889
3|001356679
4|22444466788
5|000

The third quartile of the mark distribution is equal to:

(A) 75
(B) 44
(C) 32
(D) 37.5

Only for reference. Do not copy

117. Which of the following is NOT CORRECT?

(A) The scatter plot is the basic graphical tool for investigating relationships between two continuous interval or ratio scaled variables
(B) The frequency table is useful for summarizing data from a nominal scaled variable
(C) Means and standard deviations of nominal or ordinal scaled variables are useful summary measures
(D) Pie charts don't perform well because people have difficulty in accurately quantifying angles

118. A business evaluates a proposed venture. It stands to make a profit of 10,000 with probability 3/20, to make a profit of 5,000 with probability 9/20, to break even with probability 1/4 and to lose 5,000 with probability 3/20. The expected profit is

(A) 1,500
(B) 3,000
(C) 3,250
(D) –1,500

119. In systematic sampling with $N = 20$ and $n = 5$, the number of possible samples will be

(A) 4
(B) 5
(C) 3
(D) 2

120. The sample mean is an unbiased estimator for the population mean. This means:

(A) the sample mean always equals the population mean
(B) the average sample mean, over all possible samples, equals the population mean
(C) the sample mean is always very close to the population mean
(D) the sample mean has a normal distribution

121. Age of 200 players in a hockey league follow a distribution that is skewed to the left. Suppose random samples of 15 players are taken from this population and sample mean age of players in each sample is calculated. The shape of the sampling distribution of sample mean is

(A) skewed to the right
(B) approximately normal
(C) skewed to the left
(D) difficult to determine

Only for reference. Do not copy

122. Which statement is NOT CORRECT?

(A) The sample standard deviation measures variability of our sample values
(B) A larger sample will give answers that vary less from the true value
(C) The sampling distribution describes how our estimate will vary if a new sample is taken
(D) A large sample size always gives unbiased estimators regardless of how the sample is chosen

123. In an exam, marks obtained by students in Physics, Mathematics and Statistics denoted by X, Y and Z are Normally distributed with means 50, 52 and 48 respectively and with standard deviations equal to 15, 12 and 16 respectively. The distribution of $(X + Y + Z)$ is

(A) $N(150, 43)$
(B) $N(50, 625)$
(C) $N(150, 625)$
(D) $N(50, 43)$

124. Which of the following is CORRECT?

(A) We do not need to randomize if our sample size is sufficiently large
(B) A large sample size always ensures that our sample is representative of the population
(C) In a properly chosen sample, an estimate will be less variable with a large sample size and hence more precise
(D) In random samples, the randomization ensures that we get precise and accurate estimates

125. Let $X_1, X_2, \ldots, X_n$ be independent random variables each Uniform on $[0, \theta]$ where $\theta > 0$. Then complete and sufficient statistic for θ is

(A) $\min(X_1, X_2, \ldots, X_n)$
(B) $\sum_{i=1}^{n} X_i$
(C) $\prod_{i=1}^{n} X_i$
(D) $\max(X_1, X_2, \ldots, X_n)$

Only for reference. Do not copy

126. In a shoe factory, processes A, B and C respectively manufacture 25%, 35% and 40% of the total shoes. Of their respective productions, 5%, 4% and 2% of the shoes are defective. A shoe is selected at random from the production of a particular day. If it is found to be defective, the probability that it is manufactured by the process C is

(A) $\frac{16}{69}$

(B) $\frac{25}{69}$

(C) $\frac{28}{69}$

(D) $\frac{25}{44}$

127. Let X and Y be two random variables taking values

X: 1, 2, 3, 4, 5, 6, 7, 8, 9;
Y: 1, 2, 3, 4, 5, 6, 7, 8, 9

each with equal probability. Then $P(X + Y = 10) + P(X + Y = 8)$ is

(A) $\frac{7}{66}$

(B) $\frac{10}{21}$

(C) $\frac{16}{81}$

(D) $\frac{7}{21}$

128. Let X be a random variable with mean 3 and standard deviation 2 and let Y be a random variable with mean 4 and standard deviation 1. X and Y have a correlation coefficient of –0.3. Then $\text{cov}(X, X + Y)$ is

(A) 3.8
(B) 3.4
(C) 4.8
(D) 4.2

Only for reference. Do not copy

129. The mean of 1/X when X follows F distribution with 4, 8 degrees of freedom is

(A) $\frac{4}{3}$
(B) 2
(C) $\frac{1}{2}$
(D) $\frac{3}{4}$

130. Which of the following would generally require the largest sample size?

(A) Cluster Sampling
(B) Simple Random Sampling
(C) Systematic Sampling
(D) Proportional Stratified Sampling

131. The median of exponential distribution (4) is

(A) $\ln 2$
(B) $\ln 2/16$
(C) $\ln 2/4$
(D) 4

132. A box contains 10 diskettes, of which 2 are defective. If 4 diskettes are drawn randomly one by one without replacement and X denotes the number of defective diskettes drawn, then probability of at least one defective is

(A) 0.66
(B) 0.33
(C) 0.16
(D) 0.12

133. The points of inflexion of Normal curve are given by

(A) $\mu \pm 3\sigma$
(B) $\mu \pm 2\sigma$
(C) $\pm \sigma$
(D) $\mu \pm \sigma$

Only for reference. Do not copy

134. Which of the following distributions does not possess the reproductive property?

(A) Normal distribution
(B) Poisson distribution
(C) Chi-square distribution
(D) Exponential distribution

135. The moment generating function of a random variable X is $M_X(t) = \frac{2}{5} + \frac{1}{3}e^{2t} + \frac{4}{15}e^{3t}$. The expected value of X is

(A) $\frac{22}{15}$
(B) $\frac{9}{5}$
(C) $\frac{17}{15}$
(D) $\frac{11}{5}$

136. Let $X \sim \exp(\lambda)$, then distribution of $Y = \frac{X}{2}$ is

(A) $\exp\left(\lambda^2\right)$
(B) $\exp\left(\frac{\lambda}{2}\right)$
(C) $\exp(2\lambda)$
(D) Gamma (λ)

137. If $X \sim b(n, p)$, then the distribution of $Y = (n - X)$ is

(A) $b(n,1)$
(B) $b(n, p^2)$
(C) $b(n, p)$
(D) $b(n,1-p)$

Only for reference. Do not copy

138. A necessary and sufficient condition for existence of a basic optimum feasible solution for maximization of objective function in a Linear Programming problem is that (for all j): (under usual notation).

(A) $z_i - c_j \geq 0$

(B) $z_i - c_j \leq 0$

(C) $z_i - c_j = 0$

(D) $z_i - c_j > 0$ or $z_i - c_j < 0$

139. The maximum likelihood estimates are generally:

(A) unbiased and inconsistent
(B) unbiased and consistent
(C) consistent and invariant
(D) invariant and unbiased

140. A box of 1000 items is subjected to an acceptance sampling plan that examines only 50 items. The actual fraction defective in the box is 0.02 and the sampling plan has a 0.53 probability of accepting a box of this quality. The average outgoing quality will be

(A) 0.02
(B) 0.51
(C) 0.53
(D) 0.01

141. Let $X_1, X_2, \ldots, X_n$ be independent random variables each Bernoulli with p as probability of success. Then sufficient statistic for p is

(A) $\bar{X}$

(B) $\sum_{i=1}^{n} X_i$

(C) $\prod_{i=1}^{n} X_i$

(D) $\sum_{i=1}^{n} X_i^2$

Only for reference. Do not copy

142. Rao-Blackwell theorem enables us to obtain minimum variance unbiased estimator through

(A) unbiased estimators
(B) complete statistics
(C) efficient statistics
(D) sufficient statistics

143. If T is an unbiased estimator of θ, then T^2 is

(A) unbiased estimator of θ^2
(B) biased estimator of θ^2
(C) unbiased estimator of θ^2+1
(D) biased estimator of θ^2+1

144. A sample is drawn from the population with pdf $f(x)=\frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{Gamma(\alpha)}, x\geq 0$ and first and second order sample moments about origin are 12 and 180 respectively. Then estimators of α and β obtained by method of moments are

(A) $\hat{\alpha}=5$ and $\hat{\beta}=1/3$
(B) $\hat{\alpha}=5$ and $\hat{\beta}=2/3$
(C) $\hat{\alpha}=3$ and $\hat{\beta}=5/3$
(D) $\hat{\alpha}=3$ and $\hat{\beta}=2$

145. In designing an experiment, blocking is used

(A) to reduce bias
(B) to reduce variation
(C) as a first step in randomization
(D) to control the level of the experiment

146. Data is obtained from 6 samples each consisting of 5 observations and the equality of six population means is to be tested. For the critical value of F, the degrees of freedom are

(A) 5 for numerator and 24 for denominator
(B) 30 for numerator and 24 for denominator
(C) 5 for numerator and 30 for denominator
(D) 6 for numerator and 24 for denominator

Only for reference. Do not copy

Only for reference. Do not copy

147. In terms of joint cumulative distribution function $F(x, y)$, $P(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2)$ is

(A) $F(a_1, b_1) - F(a_1, b_2) - F(a_2, b_1) - F(a_2, b_2)$
(B) $F(a_1, b_1) + F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1)$
(C) $F(a_1, b_2) - F(a_1, b_1) - F(a_1, b_2) + F(a_2, b_1)$
(D) $F(a_1, b_2) + F(a_2, b_2) + F(a_1, b_2) + F(a_2, b_1)$

148. Let U and V be two independent random variables each following Uniform distribution on (0, 1). Then distribution of $Z = \max(X, Y)$ is

(A) Beta (1, 1)
(B) Beta (2, 2)
(C) Beta (1, 2)
(D) Beta (2, 1)

149. Mean squared error of an estimator T_n of $\tau(\theta)$ is expressed as

(A) bias + $\text{var}_\theta(T_n)$
(B) $[\text{bias} + \text{var}_\theta(T_n)]^2$
(C) $(\text{bias})^2 + [\text{var}_\theta(T_n)]^2$
(D) $(\text{bias})^2 + \text{var}_\theta(T_n)$

150. A continuous random variable X has pdf as

$$f(x) = \begin{cases} \dfrac{3x^2}{\theta^3}, \text{ for } 0 \leq x \leq \theta \text{ for some } \theta > 0 \\ 0, \quad \text{otherwise} \end{cases}.$$

Then ratio of mode to median is

(A) $2^{-\frac{1}{3}}$
(B) $2^{\frac{1}{3}}$
(C) $\dfrac{\theta}{2^{\frac{1}{3}}}$
(D) θ

Only for reference. Do not copy

ANSWER KEY

Subject Name: 614 STATISTICS

SI No.	Key	SI No.	Key	SI No.	Key	SI No.	Key	SI No.	Key
1	A	31	C	61	A	91	B	121	C
2	B	32	A	62	A	92	A	122	D
3	A	33	A	63	A	93	D	123	C
4	A	34	B	64	C	94	B	124	C
5	A	35	D	65	D	95	C	125	D
6	B	36	B	66	D	96	A	126	A
7	C	37	C	67	B	97	B	127	C
8	A	38	B	68	D	98	C	128	B
9	B	39	B	69	A	99	D	129	B
10	C	40	B	70	D	100	C	130	A
11	A	41	C	71	B	101	C	131	C
12	B	42	B	72	B	102	C	132	A
13	D	43	D	73	A	103	B	133	D
14	D	44	D	74	B	104	B	134	D
15	C	45	D	75	A	105	A	135	A
16	C	46	C	76	A	106	A	136	C
17	B	47	C	77	C	107	C	137	D
18	B	48	C	78	D	108	C	138	B
19	A	49	A	79	C	109	C	139	C
20	A	50	B	80	B	110	B	140	D
21	D	51	A	81	D	111	B	141	B
22	B	52	B	82	A	112	C	142	D
23	A	53	C	83	B	113	D	143	B
24	B	54	A	84	B	114	C	144	A
25	C	55	B	85	C	115	A	145	B
26	B	56	B	86	A	116	B	146	A
27	C	57	A	87	D	117	C	147	B
28	B	58	C	88	B	118	B	148	D
29	B	59	D	89	A	119	A	149	D
30	C	60	B	90	B	120	B	150	B

Only for reference. Do not copy