

Code No.

N – 3581

Entrance Examination for Admission to the P.G. Courses in the Teaching Departments, 2022

CSS

STATISTICS/APPLIED STATISTICS AND DATA ANALYTICS

General Instructions

1. The Question Paper is having two Parts — Part 'A' Objective type (60%) & Part 'B' Descriptive type (40%).
2. Objective type questions which carry 1 mark each are to be (✓) 'tick marked' in the response sheets against the appropriate answers provided.
3. 8 questions are to be answered out of 12 questions carrying 5 marks each in Part 'B'.
4. **Negative marking** : 0.25 marks will be deducted for each wrong answer in Part 'A'.

Time : 2 Hours**Max. Marks : 100**

To be filled in by the Candidate									
Register Number	in Figures								
	in words								

PART – A

(Objective Type)

Choose appropriate answer from the options in the questions. **One mark each.**

(60 × 1 = 60 marks)

1. If a data set has an even number of observations, the median is
 - a) Equal to the mean value
 - b) Cannot be determined
 - c) Between two middle value
 - d) Average value of the two middle value arranged in an ascending order

DO NOT WRITE HERE

-
2. The degrees of freedom for performing an F-test are given by
- a) n_1, n_2 b) $n_1 + 1, n_2 + 1$ c) $n_1 - 1, n_2 - 1$ d) $n_1 + n_2 - 1$
3. In textile industry, a manufacturer is interested in the number of blemishes or flaws occurring in each 100 feet of material. The probability distribution that has the greatest chance of applying to this situation is the
- a) Normal distribution b) Binomial distribution
c) Poisson distribution d) Uniform distribution

4. A numerical description of the outcome of an experiment is called a
 - a) Descriptive statistic
 - b) Probability function
 - c) Variance
 - d) Random variable

5. Mc-nemar test is used when
 - a) Independence
 - b) Dependent
 - c) Association of attributes
 - d) Categorical

6. The angle between the minute hand of a clock and the hour hand of a clock when the time is 7 : 05 AM is
 - a) 80 degree
 - b) 100 degree
 - c) 180 degree
 - d) 90 degree

7. A city has a population of 10 Lakhs out of which 4.5 Lakhs are males. 30% of population is illiterate and if 60% of female population are literate the number of literate male is
 - a) 3,10,000
 - b) 2,20,000
 - c) 3,30,000
 - d) 3,70,000

8. The sampling distribution refers to
 - a) The distribution of various sample sizes which might be used in a given study
 - b) The distribution of the values of the items in the population
 - c) The distribution of the different possible values of the sample mean together with their respective probabilities of occurrence
 - d) The distribution of the values of the items actually selected in a given sample

9. A secretary goes to work in three routes A, B and C. Her choice of route is independent of the weather conditions. If it rains, the probabilities of arriving late, are given as 0.06, 0.15 and 0.12. If it does not rain then 0.05, 0.10 and 0.15. On a particular day she arrives late, the probability that it is a rainy day is
 - a) $\frac{41}{131}$
 - b) $\frac{1}{2}$
 - c) 0.05
 - d) $\frac{41}{100}$

10. Which of the following is related to asymptotic property of an estimator?
 - a) Unbiasedness
 - b) Consistency
 - c) Efficiency
 - d) Sufficiency

11. National census day is observed on
 a) April 1 b) June 29 c) June 8 d) February 9
12. In a sample survey, the chances of admission in M.Sc., Statistics courses and M.Sc., Data science course or both are 0.89, 0.35 and 0.29 respectively. What is the probability of owning either or both degrees?
 a) 0.82 b) 0.92 c) 0.08 d) 0.18
13. If $F(x)$ is a distribution function of a random variable X and given statements

 (1) : $F(-\infty)=0$ and $F(+\infty)=1$
 (2) : $F(-\infty)=0$ and $F(+\infty)=0$
 (3) : $a < b \rightarrow F(a) \leq F(b)$ for real a and b
 (4) : $a < b \rightarrow F(a) \geq F(b)$ for real a and b
 a) (1) and (4) are true b) (2) and (4) are true
 c) (1) and (3) are true d) (2) and (3) are true
14. The probability distribution function always satisfies the postulates
 a) Always true b) Partially true
 c) Always false d) Partially false
15. If random variable x has the probability function $f(x) = \frac{k}{x!}$ ($x = 0, 1, 2, \dots$) the value of k is _____
 a) $\frac{1}{e}$ b) e c) $1 - \frac{1}{e}$ d) None of these
16. In normal distribution, the curve of $f(x)$ is symmetric with respect to _____
 a) $x = \sigma$ b) $x = \mu$ c) $x = \frac{1}{\sigma}$ d) None of these
17. Let $\Phi(x)$ be a standard normal distribution function. Then following relation holds

 a) $\Phi(-x) = 1 - \Phi(x)$ b) $\Phi(-x) = \Phi(x)$
 c) $\Phi(-x) = \Phi(x) - 1$ d) None of these

18. Repeated trials are normally referred to
- Uniform distribution
 - Poisson distribution
 - Gamma distribution
 - Bernoulli distribution
19. If you buy one ticket in the provincial Lottery, then the probability that you will win a prize is 0.11. If you buy one ticket for five months, what is the probability that you will win at least one prize?
- 0.55
 - 0.50
 - 0.44
 - 0.45
20. Consider 8 blood donors chosen randomly from a population. The probability that the donor has type A blood is 0.40. Which of the following is correct?
- The probability of 1 or fewer donors having type A blood is about 0.11
 - The probability of 7 or more donors NOT having type A blood is 0.0087
 - The probability of exactly 5 donors having type A blood is about 0.28
 - The probability of exactly 5 donors not having type A blood is about 0.12
21. An insurance company has estimated the following cost probabilities for the next year on a particular model of car
- | | | | | |
|------|-------|---------|----------|----------|
| Cost | Rs. 0 | Rs. 500 | Rs. 1000 | Rs. 2000 |
| Prob | .60 | .50 | .13 | ??? |
- The expected cost to the insurance company is (approximately)
- Rs.155
 - Rs.595
 - Rs.875
 - Rs.495
22. The heights of students at a college are normally distributed with a mean of 175 cm and a standard deviation of 6 cm. One might expect in a sample of 1000 students that the number with heights less than 163 cm is _____
- 173
 - 22
 - 145
 - 23
23. Which of the following is not applicable to a Poisson distribution?
- It is used to compute the probability of rare events
 - Every event is independent of every other event
 - It is parameterized by the sample size and the probability that an event will occur.
 - The theoretical range for the number of events that could occur is 0,1,2,3 ...
 - In order to compute the parameter value, we need to know the standardized rate and the sample size.

24. Power test refers to _____
- The ability to detect an effect of malathion when in fact there is no effect
 - The ability to not detect an effect of malathion when in fact there is no effect
 - The ability to detect an effect of malathion when in fact there is an effect
 - The ability to not detect an effect of malathion when in fact there is an effect
25. The average time it takes for a person to experience pain relief from aspirin tablet is 25 minutes. A new ingredient is added to help speed up relief. Let μ denote the average time to obtain pain relief with the new product. An experiment is conducted to verify if the new product is better. What are the null and alternative hypothesis?
- $H_0 : \mu = 25$ vs $H_A : \mu \neq 25$
 - $H_0 : \mu = 25$ vs $H_A : \mu < 25$
 - $H_0 : \mu = 25$ vs $H_A : \mu = 25$
 - $H_0 : \mu = 25$ vs $H_A : \mu > 25$
26. In testing $H_0 : \mu = 100$ against $A : \mu \neq 100$ at the 10% level of significance, H_0 is rejected if _____
- The p-value is than 0.10
 - The value of the test statistic is in the acceptance region
 - 100 is contained in the 90% confidence interval
 - The p-value is greater than 0.10
27. What is a statistical inference?
- A decision, estimate, prediction, or generalization about the population based on information contained in a sample
 - A statement made about a sample based on the measurements in that sample
 - A set of data selected from a larger set of data to be inferred
 - A decision, estimate, prediction or generalization about sample based on information contained in a population
28. The diameter of ball bearings are known to be normally distributed with unknown mean and variance. A random sample of size 25 gave a mean 2.5 cm. The 95% confidence interval had length 4 cm. Then _____
- The sample variance is 4.86
 - The sample variance is 26.03
 - The population variance is 4.84
 - The sample variance is 23.47

29. A college student is interested in estimating the average number of showers per week. Based on a preliminary sample he believes that σ^2 is close to 2.1. How large a sample is needed if his estimate is to be within 0.3 with probability 0.95?
- a) 183 b) 90 c) 64 d) 253
30. A single stem-and-leaf plot is a useful tool because _____
- a) It includes the average and the standard deviation
- b) It shows the percentage distribution of the data values
- c) It enables us to examine the data values for the presence of trends, cycles and seasonal variation
- d) It enables us to locate the center of the data, see the overall shape of the distribution and look to marked deviations from the overall shape
31. Which one of the following statements is false?
- a) Pie charts are better than bar graphs for comparing relative sizes
- b) Data that are nominal scale are presented using frequency tables
- c) Means and standard deviation of ordinal data are meaningless
- d) The scatter-plot is the basic graphic tool for investigating relationships between two interval or ratio scaled variables
32. Which of the following is not correct?
- a) The scatterplot is the basic graphical tool for investigating relationships between two continuous intervals 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
33. Last year a small statistical consulting company paid each of its five statistical clerks \$22,000, two statistical analysts \$50,000 each, and the senior statistician/owner \$270,000. The number of employees earning less than the mean salary is _____
- a) 0 b) 4 c) 5 d) 7

34. The heights in centimeters of 5 students are
165, 175, 176, 159, 170
The sample median and sample mean are respectively
a) 170, 169 b) 170, 170 c) 169, 170 d) 176, 169
35. In measuring the center of the data from a skewed distribution, the median would be preferred over the mean for most purposes because
a) The median is the most frequent number while the mean is most likely
b) The mean may be too heavily influenced by the larger observations and this gives too high an indication of the center
c) The median is less than the mean and smaller numbers are always appropriate for the center
d) The mean measures the spread in the data
36. The frequency distribution of the amount of rainfall in December in a certain region for a period of 30 years is given below:
- | Rainfall
(in inches) | Number
of years |
|-------------------------|--------------------|
| 2.0–4.0 | 3 |
| 4.0–6.0 | 6 |
| 6.0–8.0 | 8 |
| 8.0–10.0 | 8 |
| 10.9 –12.0 | 5 |
- The mean amount of rainfall in inches is
a) 7.30 b) 7.25 c) 7.40 d) 8.40
37. What is the best reason for performing a paired experiment rather than a two-independent sample experiment?
a) It is easier to do because we need fewer experimental units and each unit receives more than one treatment
b) It allows us to remove variation in the results caused by other factors because we can compare both treatments within the same experimental unit
c) It is more accurate because it works only with the differences
d) It requires fewer assumptions because we are only interested in the difference between treatments

38. 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 case, the median is halfway between the first and the third quartiles
 - d) The median is always greater than the mean
39. An airplane is only allowed a gross passenger weight of 8000 kg. If the weights of passengers travelling by air between Toronto and Vancouver have a mean of 78 kg, and a standard deviation of 7 kg, the approximate probability that the combined weight of 100 passengers will exceed 8,000 kg is _____
- a) 0.0022 b) 0.3987 c) 0.1103 d) 0.4978
40. 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 will only vary a little from the population mean
41. In a statistical test of hypothesis, what happens to the rejection region when α , the level of significance, is reduced?
- a) The answer depends on the value of α
 - b) The rejection region is increased in size
 - c) The rejection region is reduced in size
 - d) The rejection region is unaltered
42. Since α = probability of Type I error, then $1 - \alpha$ is
- a) Probability of rejecting H_0 when H_0 is true
 - b) Probability of not rejecting H_0 when H_0 is true
 - c) Probability of not rejecting H_0 when H_A is true
 - d) Probability of rejecting H_0 when H_A is true

43. If the correlation between body weight and annual income were high and positive, we could conclude that ____
- High incomes cause people to eat more food
 - Low incomes cause people to eat less food
 - High income people tend to spend a greater proportion of their income on food than low income people, on average
 - High income people tend to be heavier than low income people, on average
44. Men tend to marry women who are slightly younger than themselves. Suppose that every man married a woman who was exactly 0.5 of a year younger than themselves. Which of the following is correct?
- The correlation is -0.5
 - The correlation is 0.5
 - The correlation is 1
 - The correlation is -1
45. For children between the ages of 18 months and 29 months, there is approximately a linear relationship between 'height' and 'age'. The relationship can be represented by:
- $$\hat{Y} = 64.93 + 0.63(x), (x), \text{ where } Y \text{ represents height (in centimeters) and } X \text{ represents age (in months).}$$
- Joseph is 22.5 months old and is 80 centimeters tall. What is Joseph's residual?
- 79.1
 - 0.9
 - -0.9
 - 56.6
46. The average time between infection with the AIDS virus and developing AIDS has been estimated to be 8 years with a standard deviation of about 2 years. Approximately what fraction of people develop AIDS within 4 years of infection?
- 5%
 - 2.5%
 - 32%
 - 16%
47. The distribution of the heights of students in a large class is roughly bell-shaped. Moreover, the average height is 68 inches, and approximately 95% of the heights are between 62 and 74 inches. Thus, the standard deviation of the height distribution is approximately equal to:
- 3
 - 2
 - 6
 - 9

48. Let X be exponentially distributed with parameter λ
- Maximum likelihood estimator of λ is the reciprocal of sample mean
 - Maximum likelihood estimator of λ is the proportional of sample mean
 - $\hat{\lambda} = 1/\bar{X}$
 - $\hat{\lambda} = \bar{X}$
49. In a random sample, 136 of 400 persons given a flu vaccine experienced some discomfort. Construct a 95% confidence interval for the true proportion of persons who will experience some discomfort from the vaccine.
- [0.204, 0.306]
 - [0.29, 0.39]
 - [1, 2]
 - [0.6, 0.9]
50. A monotone sequence $\{a_n\}_{n=1}^{\infty}$, is convergent _____
- It is bounded
 - It is unbounded
 - It is decreasing
 - It is increasing
51. A: Every convergent sequence is a Cauchy sequence
B: Every Cauchy sequence is a convergent sequence
- A and B both are false
 - A is true
 - B is true
 - A and B both true
52. A function $f(x) = |x|$ is _____
- Continuous and differentiable at $x = 0$
 - Continuous but not differentiable at $x = 0$
 - Discontinuous and differentiable at $x = 0$
 - None of these
53. Given a real number a and b such that $a \leq b + \epsilon, \forall \epsilon > 0$
- $a = b$
 - $a \geq b$
 - $a \leq b$
 - $ab = \epsilon$

54. For closed interval $[a, b]$, where $a < b, x \in [a, b]$ we have _____
 - a) $a \leq x \leq b$
 - b) $a < x \leq b$
 - c) $a \leq x < b$
 - d) $a < x < b$
55. A: sequence $\{X_n\}$ is convergent
B: sequence $\{X_n\}$ is bounded then
 - a) $A \Rightarrow B$
 - b) $B = A$
 - c) $A \Leftrightarrow B$
 - d) None of these
56. The matrix $\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$ is a
 - a) Hermitian matrix
 - b) Skew-Hermitian
 - c) Skew-Symmetric
 - d) Symmetric
57. A real quadratic form $X^T A X$ is positive definite, if _____
 - a) All eigen values of $A > 0$
 - b) All eigen values of $A < 0$
 - c) All eigen values of $A = 0$
 - d) None of these
58. If A and B are n-order square matrix and AB and BA have characteristic roots λ and β , then
 - a) $\lambda \neq \beta$
 - b) $\lambda > \beta$
 - c) $\lambda < \beta$
 - d) $\lambda = \beta$
59. Randomization in an experiment helps to eliminate
 - a) Systematic influences
 - b) dependence among observations
 - c) Human biases
 - d) all the above
60. If a statistical model fails to provide an exact test about the treatment effects, the model is classified as:
 - a) Unspecified model
 - b) Complete specified model
 - c) Incomplete specified model
 - d) None of the above

ANSWER SHEET — PART – A

1	A	B	C	D	E
2	A	B	C	D	E
3	A	B	C	D	E
4	A	B	C	D	E
5	A	B	C	D	E
6	A	B	C	D	E
7	A	B	C	D	E
8	A	B	C	D	E
9	A	B	C	D	E
10	A	B	C	D	E
11	A	B	C	D	E
12	A	B	C	D	E
13	A	B	C	D	E
14	A	B	C	D	E
15	A	B	C	D	E
16	A	B	C	D	E
17	A	B	C	D	E
18	A	B	C	D	E
19	A	B	C	D	E
20	A	B	C	D	E

21	A	B	C	D	E
22	A	B	C	D	E
23	A	B	C	D	E
24	A	B	C	D	E
25	A	B	C	D	E
26	A	B	C	D	E
27	A	B	C	D	E
28	A	B	C	D	E
29	A	B	C	D	E
30	A	B	C	D	E
31	A	B	C	D	E
32	A	B	C	D	E
33	A	B	C	D	E
34	A	B	C	D	E
35	A	B	C	D	E
36	A	B	C	D	E
37	A	B	C	D	E
38	A	B	C	D	E
39	A	B	C	D	E
40	A	B	C	D	E

41	A	B	C	D	E
42	A	B	C	D	E
43	A	B	C	D	E
44	A	B	C	D	E
45	A	B	C	D	E
46	A	B	C	D	E
47	A	B	C	D	E
48	A	B	C	D	E
49	A	B	C	D	E
50	A	B	C	D	E
51	A	B	C	D	E
52	A	B	C	D	E
53	A	B	C	D	E
54	A	B	C	D	E
55	A	B	C	D	E
56	A	B	C	D	E
57	A	B	C	D	E
58	A	B	C	D	E
59	A	B	C	D	E
60	A	B	C	D	E

STATISTICS/APPLIED STATISTICS AND DATA ANALYTICS

PART – B

(Descriptive Type)

Answer **any eight** questions. Each question carries Five marks. **(8 × 5 = 40 Marks)**

1. State the mathematical model used in analysis of variance in a two-way classification method.
2. Write a short note on circular systematic sampling.
3. State and prove chebychev's inequality.
4. Write a note of components of time series and its uses.
5. 'Index numbers are economic barometers' — Explain.
6. The following table gives the number of workers employed in a small industry during the years 2000—2008.

Year :	2000	2001	2002	2003	2004
No. of workers :	430	470	450	460	470
Year :	2005	2006	2007	2008	
No. of workers :	470	500	590	480	

Calculate the four yearly moving average.

7. The regression equation calculated from a given set of observations are $x = -0.2y + 4.2$ and $y = 0.8x + 8.4$. Calculate :
 - (a) x and y ,
 - (b) find correlation co-efficient and
 - (c) find the estimated value of y when $x = 4$.

8. In an intelligent test administrated to 1000 students the mean score was 42 and standard deviation is 24. Find (a) the number of student exceeding a score of 500. (b) The number of student lying between 30 and 54. (c) The value of score exceeded by the top 100 students.
9. State and prove crammer-rao inequality.
10. A survey of male children in 128 families each having 5 children gave the following data:
- | | | | | | | | |
|---------------------------|---|----|----|----|----|----|-------|
| Number of male children : | 0 | 1 | 2 | 3 | 4 | 5 | Total |
| Number of families : | 9 | 17 | 26 | 39 | 22 | 12 | 125 |
- a) Fit a binomial distribution for the data,
- b) When $n = 10$, and $p = 0.35$ find μ_2, μ_3 and μ_4 .
11. If $A_n = [1/n, 2]$, Prove that the limit exists or not.
12. How the growth of a population is measured? Mention some important growth curves used for measuring the growth of the population.
