| Code No. | J – 2279 |
|---|---|

**Entrance Examination for Admission to the P.G. Courses in the Teaching Departments, 2020**

**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. After finding the average of 35 scores, a student carelessly included the average with the 35 scores and found the average of these 36 numbers. The ratio of the second average to the true average will be

   a) 1 : 1 b) 35 : 36

   c) 36 : 35 d) 2 : 1

---

2. Given two numbers $x_1$ and $x_2$ the variance of these numbers is proportional to

   a) $(x_1 + x_2)^2$ b) $(x_1 - x_2)^2$

   c) $x_1^2 + x_2^2$ d) $x_1^2 - x_2^2$

3. A man travels from Kottayam to Chengannur at an average speed of 40 km/hour and returns from Chengannur to Kottayam by increasing his speed by 50%. Then the average speed for the entire trip is

   a) 50 b) 48

   c) 40 d) 55

4. The sum of squares of 10 observations from the mean is 360 and the mean is 60. Then the coefficient of variation is

a) 10% b) 20%

c) 15% d) 50%

5. The deviations of the first 3 observations of a series of 4 observations from their mean "*m*" are 4, –2 and 3. Then which of the following is true regarding the $4^{th}$ observation.

a) $x_4 = m$ b) $x_4 < m$

c) $x_4 = 3m$ d) $x_4 > m$

6. The average monthly salary workers in a factory is Rs. 206. If the average monthly salary of males and females are Rs. 210 and Rs. 190 respectively, the percentage of females employed in the factory is

a) 10 b) 50

c) 30 d) 20

7. The A.M. of 52 successive positive integers beginning with 2 is

a) 27 b) 28

c) 27.25 d) 27.5

8. The median length of 50 items is 52 cm. It was later found that the largest observation was wrongly reported as 100 instead of 110 cm. The correct median is

a) 62 b) 52

c) 42 d) Not determinable

9. If a constant "*m*" is added to every item of a series, then which one of the following measures of dispersion will change?

a) Mean Deviation about Mean b) Range

c) Standard Deviation d) Coefficient of variation

10. Suppose 3 observations on the ordered pair $(x, y)$ yielded $(0, -1)$, $(-1, 0)$ and $(1, 1)$. The least square estimate of $\beta$ in fitting a line of the form $y = \beta x$ is

a) $-\frac{1}{2}$ b) $\frac{1}{2}$

c) $\frac{2}{3}$ d) 1

11. In a bivariate study the correlation coefficient $r = 0.8$. If all the observations on *X* and *Y* are reduced by 50% then the correlation coefficient will be

a) 0.3 b) 0.4

c) 0.8 d) 0

12. If the lines of regression of *X* on *Y* and *Y* on *X* are respectively $a_1 X + b_1 Y = C_1$ and $a_2 X + b_2 Y = C_2$ then

a) $\frac{a_1 b_2}{a_2 b_1} \geq 1$ b) $\frac{a_1 b_2}{a_2 b_1} < 1$

c) $\frac{a_1 b_2}{a_2 b_1} = 0$ d) None of the above

13. All the ordinary correlation coefficients in a set of 3 variables $X_1$, $X_2$, $X_3$ are equal to *r*. Then the value of partial correlation coefficient $r_{12.3}$ is

a) $\frac{1}{1+r}$ b) $\frac{2}{1+r}$

c) $\frac{1-r}{1+r}$ d) $\frac{r}{1+r}$

14. Which of the following is true?

a) Fisher's Ideal Index satisfies only the Time Reversal Test

b) All Index numbers satisfy the Commodity Reversal Test

c) Laspeyre's Index satisfy Factor Reversal Test

d) Paasche's Index satisfy Time Reversal Test

15. If Laspeyre's Price Index is 208 and Paasche's Index is 52 then Fisher's Index is

a) 82 b) 74

c) 104 d) 120

16. With which characteristic movement of a time series can the increasing demand for two-wheelers be associated?

a) Secular Trend
b) Seasonal Variation
c) Cyclical Variation
d) Irregular Variation

17. $X \sim U[0, 1]$. Let $Y = F(X)$ where $F(x)$ is the distribution function $X$. Then the distribution of $Y$ is

a) Uniform
b) Cauchy
c) Normal
d) Exponential

18. The series $\sum \frac{1}{n^p}$ converges if

a) $p \leq 1$
b) $p > 1$
c) $p = 1$
d) $p \geq 1$

19. Which of the following statements is not true?

(i) Every subset of a countable set is countable
(ii) Set of all rational numbers is countable
(iii) The set of irrational numbers is uncountable
(iv) The set of real numbers in $[0\ 1]$ is countable

a) only (iii)
b) (i) and (iii)
c) only (iv)
d) (iii) and (iv)

20. Which of the following statements is true?

(i) A sequence can have several limits
(ii) A convergent sequence is bounded
(iii) Any bounded sequence converges
(iv) A Cauchy sequence does not converge always

a) only (iii)
b) only (ii)
c) only (i)
d) only (iv)

21. The series $1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + - \cdots$

a) converges
b) diverges to $\infty$
c) diverges to $-\infty$
d) oscillates

22. Let the probability that a child is male be $\frac{1}{2}$. Consider a family with 3 children and the following events :

E : There are children of both sexes

F : There is atleast one girl

Which of the following statements is correct?

a) E and F are mutually exclusive
b) E and F are independent
c) F is a subset of E
d) E is subset of F

23. X has Poisson distribution with $P[X \geq 1] = \frac{1}{2}$. Then $V(X)$ will be

a) 4
b) $\log 2$
c) 5
d) $\log 3$

24. A missile protection system has 5 radar sets operating independently each with a probability 0.9 of detecting a missile entering a zone that is covered by all of the units. Then the probability that at least one detects the missile will be

a) 0.5
b) 0.3
c) $(0.1)^5$
d) $1 - (0.1)^5$

25. There are 3 children in a family. What is the probability that there are at least 2 boys given that eldest is a boy?

a) 1/4
b) 1/2
c) 1/3
d) 3/4

26. A statistician is told that he can toss an unbiased coin *n* times where *n* is even and if he gets exactly 3 heads he will be given a prize. Then the best choice for *n* is

a) 4
b) 8
c) 6
d) 10

27. Distribution of the sample mean is the same as that of each random sample unit $\Rightarrow$ that the distribution can be

a) Cauchy
b) Normal
c) Discrete distribution taking non –ve integers
d) Exponential

28. Let $X$ be a Binomial random variable with $E[X]=Var(X)$ . Which of the following statements is true.

a) Mean is twice the mode

b) The distribution is symmetric

c) The parameter '$n$' is necessarily even

d) No Binomial distribution with the above property exists

29. $X_1, X_2, \cdots, X_n$ are iid random variables following Poisson distribution with parameter $\lambda=1$ and $Y=X_1+X_2+\cdots+X_n$. What is $P(Y \le n)$ as $n \to \infty$?

a) 0.75 b) 0

c) 0.5 d) 0.05

30. A proportion 0.49 of all children in a large population are males. From all families with 3 children each, a random sample of 1000 families is selected. The expected number of families which have at least one female child is

a) 877 b) 118

c) 133 d) 882

31. A random variable has p.d.f. $f(x)=k\sqrt{x(1-x)}$ , $0 \le x \le 1$. Then $E[X]$ is

a) 1/2 b) 1/3

c) 1/4 d) 5/12

32. An electronic equipment is known to have its lifetime (in years) distributed with density 0.1 $e^{-0.1x}, x>0$. At age 4 years when the equipment is still working its expected further life in years is

a) 1 b) 5

c) 12 d) 10

33. $X_1$ and $X_2$ are independent exponential random variables with mean $\lambda$. The distribution of $\min(X_1, X_2)$ is

a) exponential with mean $\lambda/2$ b) exponential with mean $2\lambda$

c) exponential with mean $\lambda$ d) not exponential

34. Anil, Binil and Sunil go for an outing. They agree to toss an unbiased coin each until they get an odd man (one getting an outcome different from others) and the odd man has to pay for their tea. What is the probability that Anil will be the odd man at the first trial itself?

a) 1/2 b) 1/4

c) 1/3 d) 1

35. A man has 10 keys on a chain only one of which opens the door to his apartment. As the man is drunk he is not able to distinguish one key from another. He tries the keys at random. He tries one key and if it fails to open he discards it and chooses one of the remaining 9 keys and so on. What is the probability that the door opens with the second key he tries?

a) 1/2 b) 1/9

c) 1/10 d) 2/10

36. $f(x, y) = e^{-x-y}$ $x > 0, y > 0$. What is $E[X + Y]$?

a) 2.5 b) 3

c) 2 d) 4

37. If $X$ and $Y$ are independent $N(0,1)$ random variables, what is $P[XY > 0]$?

a) 0 b) 0.4

c) 0.5 d) 1

38. Let $X_1, X_2, X_3$ be independent and identically distributed random variables with density

$$f(x) = 3x^2 \quad 0 \le x \le 1$$

Let $Y = Max[X_1, X_2, X_3]$. What is $P[Y > .5]$?

a) 3/4 b) 411/512

c) 1/512 d) 511/512

39. Let $X_1, X_2, \cdots X_n$ be a random sample from the Cauchy distribution with p.d.f. given by

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

The sample median as an estimator of $\theta$ is

a) unbiased and consistent
b) unbiased and inconsistent
c) biased and consistent
d) biased and inconsistent

40. Let $X_1, X_2, \cdots X_n$ be a random sample form $U\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]$ distribution. Define $T_1 = \overline{X}$ and $T_2 = \frac{X_1 + X_n}{2}$. Then

a) Both $T_1$ and $T_2$ are unbiased for $\theta$
b) $T_1$ is unbiased for $\theta$ but $T_2$ is not
c) $T_2$ is unbiased for $\theta$ but $T_1$ is not
d) Both $T_1$ and $T_2$ are biased for $\theta$

41. The M.L.E. of $p$ based on single observation $X$ from $B(1, p)$ with parameter space restricted to $\left[\frac{1}{3} \quad \frac{2}{3}\right]$ is

a) $\frac{X+1}{3}$
b) $\frac{2X+1}{3}$
c) $\frac{X-1}{3}$
d) $\frac{2X-1}{3}$

42. A random sample of 16 animals showed an average increase in weight of 10 kg when a new diet was administered. Assuming that the increase in weight follows $N(\mu, \sigma^2)$ with $\sigma^2 = 4$, then 95% confidence interval for the increase in weight due to the new diet is

a) [9.05 10.95]
b) [8.51 11.49]
c) [9.07 10.93]
d) [9.02 10.98]

43. The maximum likelihood estimator of the parameter *b* in the density function

$$f(x; b) = \frac{1}{2} e^{-|x-b|} \quad -\infty < x < \infty$$

based on a sample $X_1, X_2, \cdots, X_n$ is

a) maximum of $(X_1, X_2, \cdots, X_n)$ b) minimum of $(X_1, X_2, \cdots, X_n)$

c) mean of $(X_1, X_2, \cdots, X_n)$ d) median of $(X_1, X_2, \cdots, X_n)$

44. In a city, the total number *N* of taxis is not known but all taxis are numbered serially from 1 to *N*. To estimate *N*, a random sample of taxis of size *n* is drawn and their serial numbers are noted down as $X_1, X_2, \cdots X_n$. If $\overline{X}$ and *M* denote the sample mean and Max $(X_1, X_2, \cdots X_n)$ respectively, estimator of *N* obtained by method of moments will be

a) $\overline{X}$ b) $M$

c) $2\overline{X} - 1$ d) $2M - 1$

45. A random sample from $U(-\theta, \theta)$ gives observations 1.5, 0, –2.0, 3.3, –4.2, –1.0. The MLE of $\theta$ is

a) 3.3 b) 4.2

c) 3.0 d) 3.5

46. If the lifetime of television tubes is assumed to be normally distributed with mean $\mu$ and s.d. 100 hours, then what is the minimum sample size we should take to be 95% confident that the error in the estimation of mean life by the sample mean will not exceed 20 hours?

a) 68 b) 97

c) 166 d) 180

47. N.P. lemma gives a general method for finding a best (most powerful) test when

a) $H_0$ is a simple and $H_1$ is a composite hypothesis

b) $H_0$ and $H_1$ are any hypothesis (simple or composite)

c) $H_0$ and $H_1$ are both simple hypothesis

d) $H_0$ and $H_1$ are both composite hypothesis

48. An instructor gives a short quiz involving 10 true-false questions. To test the hypothesis that the students are guessing, the instructor adopts the following rule. If 7 or more answers are correct, the students is not guessing. If less than 7 answers are correct, the student is guessing. Then the probability of type I error is

a) 9/64 b) 11/64

c) 1/64 d) 3/64

49. For testing whether the weights of boys and girls of age 10 years differ significantly a researcher is applying two sample *t* test. He took a random sample of 10 boys and 15 girls and obtained their weights. Using the values of sample means and sample variances the value of the *t* statistic was obtained as 1.6. Later the researcher found that the balance on which the weights were recorded was giving the weight with 1 kg more than correct weight. Then the correct value of the *t* statistic will be

a) More than 1.6 b) Less than 1.6

c) Equal to 1.6 d) Cannot be determined

50. In a Latin square design with four treatments and with one missing plot, what is the number of error degrees of freedom?

a) 5 b) 4

c) 6 d) 15

51. In a completely randomized design there are 5 treatments A, B, C, D and E. The treatments A, B, C, D are replicated 3, 4, 5 and 6 times respectively. If in the ANOVA table the degrees of freedom for error is 19, then E is replicated

a) 4 times b) 5 times

c) 6 times d) 7 times

52. In a CRD with 3 treatments every treatment is replicated 4 times. The null hypothesis $H_0 : T_1 \leq T_2$ versus $H_1 : T_1 > T_2$ is tested by using

a) *F* test with (1, 11) degrees of freedom

b) *t* test with 6 degrees of freedom

c) *t* test with 9 degrees of freedom

d) chi-square test with 1 degree of freedom

53. Consider, the following statements : In stratified random sampling

(i) The larger the size of the stratum, the larger the size of the sample to be selected therefrom

(ii) The larger the variability within a stratum, the larger should be the size of the sample from the stratum

(iii) The cheaper the cost per sampling unit in a stratum, the larger the sample from that stratum

Which of these statements is/are correct?

a) (i) and (ii)

b) (i), (ii) and (iii)

c) (i) and (iii)

d) (ii) only

54. Which one of the following indicate sampling error?

a) The errors which arise because a researcher observes only a part of the population in sample survey

b) The errors due to non-response of the sampled respondent

c) The errors occurring during the sampling process

d) The errors occurring in the stage of taking observations from sampled units

55. Which one of the following is true for a census?

a) Non-sampling error is zero

b) Sampling error is zero

c) Both sampling and non-sampling errors are non-zero

d) None of the above

56. The following statements are given for simple random sampling without replacement.

(i) Everyone of the $\binom{N}{n}$ possible samples of size $n$ from $N$ has the same probability

(ii) The probability of drawing a specified unit at the $r^{th}$ draw is equal to the probability of drawing it at the first draw

(iii) The probability of including a specified unit in a sample of size $n$ is $\frac{n}{N}$

(iv) $V(\bar{y}) = \frac{\sigma^2}{n}$ where $\bar{y}$ is the sample mean and $\sigma^2$ is the population variance.

Which of the above statements are correct?

a) (i), (ii) and (iv) b) (i), (iii) and (iv)

c) (ii), (iii) and (iv) d) (i), (ii) and (iii)

57. If a stratified sample of size 45 is to be selected by Neymann allocation from a population with $N_1 = 150$, $N_2 = 350$, $S_1^2 = 4$, $S_2^2 = 9$, then the number of units to be selected from the first stratum is

a) 20 b) 35

c) 75 d) 10

58. $X_1, X_2, \cdots, X_{10}$ are i.i.d. $N$ [0, 1] random variables. Then $E[X_2^2 + X_4^2 + X_6^2 + X_8^2 + X_{10}^2]$ is

a) 4 b) 3

c) 5 d) 10

59. From a population of size 30, a sample of size 3 was drawn. Then total number of possible samples to be drawn from the population by using SRSWR will be greater than that drawn by using SRSWOR by

a) 22940 b) 2000

c) 3500 d) 27000

60. In a test of hypothesis if the critical region is the null set $\phi$ then the values of $\alpha$ and $\beta$ will be

a) $\alpha = 0, \beta = 0$ b) $\alpha = 0, \beta = 1$

c) $\alpha = 1, \beta = 0$ d) $\alpha = 1, \beta = 1$

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 |

## APPLIED STATISTICS AND DATA ANALYTICS

PART – B

(Descriptive Type)

Answer **any eight** questions. **(8 × 5 = 40 Marks)**

1. Establish the relationship between raw moments and central moments.

2. The probability of a male child is 0.5 and the sex of children at successive births are independent. A couple continues to have children until they have at least one son and one daughter in the family. Let $N$ denote the number of children to such a couple. Then what is $E(N)$?

3. Give the analysis of one-way classification.

4. There are 3 coins identical appearance and one of which is ideal and others biased with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ for a head. One coin is chosen at random and tossed twice. If head appears both times what is the probability that the ideal coin was chosen?

5. If $A$ and $B$ are independent events show that

   (a) $A$ and $B^C$

   (b) $A^C$ and $B$

   (c) $A^C$ and $B^C$ are also independent.

6. Derive the recurrence relation for the central moments of a Poisson distribution.

7. Show that the series $1+\frac{1}{2!}+\frac{1}{3!}+\cdots$ converges.

8. Define hyper-geometric distribution. Obtain its mean and variance.

9. Prove that for the normal distribution the Q.D., M.D. about mean and S.D. are approximately in the ratio 10 : 12 : 15.

10. Establish the lack of memory property for exponential distribution.

11. $X$ and $Y$ have joint p.d.f.

$$f(x, y)=2,\ 0<x<1,\ 0<y<x.$$

Find the marginal and conditional distributions of $X$ and $Y$. Examine whether $X$ and $Y$ are independent.

12. Explain how a 100 (1– $\alpha$)% confidence interval can be constructed for the parameter $\theta$ in an exponential distribution.

———————