

**FIRST SEMESTER (CBCSS) DEGREE EXAMINATION, NOVEMBER 2020**

Statistics

STA 1C 01—INTRODUCTORY STATISTICS

(2019 Admissions)

(Multiple Choice Questions for SDE Candidates)

**Time : 15 Minutes****Total No. of Questions : 15****Maximum : 15 Marks****INSTRUCTIONS TO THE CANDIDATE**

1. This Question Paper carries Multiple Choice Questions from 1 to 15.
2. The candidate should check that the question paper supplied to him/her contains all the 20 questions in serial order.
3. Each question is provided with choices (A), (B), (C) and (D) having one correct answer. Choose the correct answer and enter it in the main answer-book.
4. The MCQ question paper will be supplied after the completion of the descriptive examination.

STA 1C 01—INTRODUCTORY STATISTICS  
(Multiple Choice Questions for SDE Candidates)

1. A study based on complete enumeration of data is known as :
- (A) Sample survey. (B) Pilot survey.  
(C) Census survey. (D) None of the above.
2. Average calculated in which all the items are not equally important is called \_\_\_\_\_.
- (A) Simple average. (B) Weighted average.  
(C) Combined arithmetic mean. (D) None of these.
3. What is the simple arithmetic mean of 15, 0, 36, 0 and 9 ?
- (A) 20. (B) 15.  
(C) 10. (D) 60.
4. The measure of central tendency defined as the middle most value of the variable, when the values are arranged in order of magnitude is :
- (A) Median. (B) Mode.  
(C) Arithmetic Mean. (D) Harmonic Mean.
5. Median =  $1 + c(N/2 - m)/f$  where 'm' is :
- (A) Cumulative frequency of the median class.  
(B) Frequency of the median class.  
(C) Cumulative frequency of the class preceding the median class.  
(D) Frequency of the class preceding the median class.
6. Which of the following measures of central tendencies need not be possessed by all distributions ?
- (A) Geometric Mean. (B) Harmonic Mean.  
(C) Median. (D) Mode.
7. What is the geometric mean of 2,4 and 8 ?
- (A) 4. (B) 6.  
(C) 5. (D) 7.

15. Phenomenon of Statistical Regularity is observed when :
- (A) Number of trials of a random experiment increases.
  - (B) Number of trials of a random experiment is kept minimal.
  - (C) Relative frequencies approach divergent values.
  - (D) None of these.

CHMK LIBRARY UNIVERSITY OF CALICUT

8. Mean marks of 150 students were found to be 50. Later it was found that a score of 87 was misread as 78. What is the correct mean ?
- (A) 49.9. (B) 59.  
(C) 50.09. (D) None of these.
9. \_\_\_\_\_ is used when only a rough measure of dispersion is required.
- (A) Mean Deviation. (B) Standard Deviation.  
(C) Quartile Deviation. (D) Range.
10. \_\_\_\_\_ is used to compare the consistency of 2 or more sets of data.
- (A) Co-efficient of Variation. (B) Co-efficient of Correlation.  
(C) Co-efficient of Kurtosis. (D) Co-efficient of Skewness.
11. \_\_\_\_\_ analysis deals with the intensity of relationship among variables.
- (A) Regression. (B) Correlation.  
(C) Both (A) and (B). (D) None of these.
12. Which among the following is the possible value of Co-efficient of Correlation between 2 sets of values ?
- (A) - 3. (B) - 1.9.  
(C) - 0.5. (D) 2.
13. Which of the following is not a characteristic of a random experiment ?
- (A) Number of outcomes is 2 or more.  
(B) All outcomes are not known in advance.  
(C) Outcome obtained in a particular trial is not known in advance.  
(D) Experiment can be repeated under identical conditions.
14. In a random experiment of rolling a die and observing the number shown up, let A be the event "odd number showing up". Then  $A^C =$  :
- (A) {1, 2, 3, 6}. (B) {1}.  
(C) {1, 3, 5}. (D) {2, 6}.

## FIRST SEMESTER (CBCSS) DEGREE EXAMINATION, NOVEMBER 2020

Statistics

STA 1C 01—INTRODUCTORY STATISTICS

(2019 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of calculator and Statistical table are permitted.***Section A (Short Answer Type Questions)***Answer at least eight questions.**Each question carries 3 marks.**All questions can be attended.**Overall Ceiling 24.*

1. What are the major divisions consisting of the statistical wing of MoSPI ?
2. Distinguish qualitative and quantitative data.
3. Name any four graphical methods for representing data.
4. If the mean of  $x_i$  values are 5, find the mean of  $u_i$  values where,  $u_i = \frac{x_i}{2} - 3$ .
5. Find the mean deviation about mean of the observations, 3, 6, 10, 12, 14 and 15.
6. Define outliers in a box plot.
7. Find mean and variance of a set of values with first two moments about 5 are 2 and 32.
8. State the principle of least squares in curve fitting.
9. Write any two properties of regression co-efficients.
10. Define cyclical and irregular variation in a time series data.
11. What is Marshal-Edgeworth formula for finding index number ? What are the terms involved ?
12. Differentiate simple and weighted index numbers.

(8 × 3 = 24 marks)

Turn over

## Section B (Short Essays/Paragraph Type Questions)

Answer at least five questions.

Each question carries 5 marks.

All questions can be attended.

Overall Ceiling 25.

13. Explain the role of CSO in Indian Statistical System.
14. Explain the points to be considered while making a questionnaire.
15. Explain skewness and measures of skewness based on quartiles and moments.
16. Explain the method of fitting a parabola  $y = ax^2 + bx + c$  using the observations  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  on  $x$  and  $y$ .
17. Calculate the co-efficient of correlation between  $x$  and  $y$  using the following data :
 

$x$	21	23	30	54	57	58	72	78	87	90
$y$	60	71	72	89	110	84	100	92	113	135
18. For two variables  $x$  and  $y$ , show that  $-1 \leq r_{xy} \leq 1$ , where  $r_{xy}$  is Pearson's co-efficient of correlation.
19. Explain the method of semi average for finding the secular trend in a time series data.

(5 × 5 = 25 marks)

## Section C (Essay Type Questions)

Answer any one question.

The question carries 11 marks.

20. Define : (i) Quartile deviation (ii) Co-efficient of variation. Calculate the quartile deviation and co-efficient of variation for the following data :
 

Class	:	5-15	15-25	25-35	35-45	45-55	55-65	65-75
Frequency	:	4	11	18	24	14	6	3
21. (i) Explain the uses of index numbers.
- (ii) Calculate Fisher's index number for the following data :

Items	Quantity		Price	
	2012	2018	2012	2018
A	11	18	22	26
B	9	12	18	21
C	13	14	7	10
D	17	20	11	16
E	12	18	12	18

(1 × 11 = 11 marks)

## FIRST SEMESTER (CBCSS—UG) DEGREE EXAMINATION, NOVEMBER 2021

Statistics

STA 1C 03– DESCRIPTIVE STATISTICS

(2019–2020 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of calculator and Statistical table are permitted.***Part A (Short Answer Type Questions)***Each question carries 2 marks.**Maximum marks that can be scored from this part is 20.*

1. Define secondary data. Mention any two sources of secondary data.
2. Define a frequency distribution.
3. Name any four different types of bar diagrams.
4. Define sampling frame.
5. Find the sum of 12 observations whose arithmetic mean is given as 15.
6. Define percentiles.
7. Define median. Write one of its merits and demerits.
8. Three quartiles of a set of observations are 4, 13 and 22 respectively. Calculate the quartile deviation of the set.
9. Establish that the variance of a set of observations is always a non-negative quantity.
10. For a set of observations  $x_1, x_2, \dots, x_n$ , show that the first raw moment is the mean.
11. When a set of observations is said to be skewed ?
12. Coefficient of kurtosis based on moments for a set of observations is calculated as 4, where the standard deviation is 6. Find the fourth central moment of the observations.

**Part B (Short Essay/Paragraph Type Questions)***Each question carries 5 marks.**Maximum marks that can be scored from this part is 30.*

13. Write any five steps to be considered while making a statistical investigation.
14. Define sampling and census. Write any three advantages of sampling over census.

Turn over

15. Calculate 4<sup>th</sup> decile of the following data :

Class	0-20	20-40	40-60	60-80	80-100
Frequency	18	22	28	18	14

16. Average daily revenue of 19 shops from a certain street is calculated as Rs. 2,560. A new shop is opened in the same street and recorded daily revenue of Rs. 1,800. What is the average daily revenue when all these 20 shops are taken together ?
17. Calculate mean deviation from median for the following set of 8 observations.  
24, 48, 34, 28, 42, 38, 30, 32.
18. Given the first three raw moments of a set of data as 6, 40 and 300. Calculate the coefficient of skeweness based on moments.
19. Explain : (i) Pearson's and (ii) Bowley's coefficients of skewness.

**Part C (Essay type Questions)**

*Each question carries 10 marks.*

*Maximum marks that can be scored from this part is 10.*

20. Define random and non-random sampling methods. Explain any two methods of random and non-random sampling.
21. Explain Lorenz curve. Make a rough sketch of the Lorenz curve for the following data :

Class	0-10	10-20	20-30	30-40	40-50	50-60
Frèquency :	6	10	24	30	16	14

FIRST SEMESTER (CBCSS—UG) DEGREE EXAMINATION  
NOVEMBER 2021

Statistics

STA 1C 02—DESCRIPTIVE STATISTICS

(2019—2020 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of calculator and Statistical table are permitted.*

**Part A (Short answer type Questions)**

*Each question carries 2 marks.*

*Maximum marks that can be scored from this part is 20.*

1. Define population and sample.
2. Define relative frequency of a class.
3. Define classification of data.
4. What is a primary data ?
5. Calculate the degree measure for marking the area of a component with frequency 24 out of the total frequency 120 in a pie diagram.
6. Find the geometric mean of the values 2, 9 and 12.
7. Arithmetic mean of a set of 10 observations is 24. If one of the observations 33 is removed from the set. What will be the revised arithmetic mean of the set ?
8. Mention any *two* limitations of AM as a measure of average.
9. Define Range.
10. Calculate the variance of the values 1, 2, 3, 4, 5, 6 and 7
11. Define relative measures of dispersion.
12. Find the SD of a set of data, if moment measure of skewness, and third central moment of a set of 10 observations are 0.2, 16 respectively.

Turn over

**Part B (Short Essay/Paragraph Type Questions)***Each question carries 5 marks.**Maximum marks that can be scored from this part is 30.*

13. Explain the steps involved in framing a frequency distribution for a given raw data.

14. Explain (i) Histogram ; and (ii) Frequency polygon.

15. Construct a less than ogive curve for the following frequency distribution :

Class	0-10	10-20	20-30	30-40	40-50
Frequency	6	11	16	12	5

16. Find first and third quartiles for the following data :

Class	Below 20	20-40	40-60	60-80	Above 80
Frequency	12	19	26	16	7

17. Calculate the mean deviation about the median of the following data on the age of patients admitted with a type of viral infection :

Age	0-10	10-20	20-30	30-40	40-50	50-60
Patients :	14	26	20	16	14	10

18. Calculate the coefficient of variation for the following frequency distribution :

Class	0-10	10-20	20-30	30-40	40-50
Frequency	6	11	16	12	5

19. Explain kurtosis and percentile measure of kurtosis.

**Part C (Essay Type Questions)***Each question carries 10 marks.**Maximum marks that can be scored from this part is 10.*

20. (i) Explain various steps involved in planning of a survey.

- (ii) Explain the different types of bar diagrams for representing data.

21. Define Skewness. Calculate the Pearson's co-efficient of skewness for the following data :

Class	0-10	10-20	20-30	30-40	40-50
Frequency :	4	7	11	6	2

**FIRST SEMESTER (CBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STA 1C 01—INTRODUCTORY STATISTICS

(2019—2020 Admissions)

(Multiple Choice Questions for SDE Candidates)

**Time : 15 Minutes**

**Total No. of Questions : 15**

**Maximum : 15 Marks**

**INSTRUCTIONS TO THE CANDIDATE**

1. This Question Paper carries Multiple Choice Questions from 1 to 15.
2. The candidate should check that the question paper supplied to him/her contains all the 15 questions in serial order.
3. Each question is provided with choices (A), (B), (C) and (D) having one correct answer. Choose the correct answer and enter it in the main answer-book.
4. The MCQ question paper will be supplied after the completion of the descriptive examination.

## STA 1C 01—INTRODUCTORY STATISTICS

(Multiple Choice Questions for SDE Candidates)

1. A study based on complete enumeration of data is known as :
- (A) Sample survey. (B) Pilot survey.  
(C) Census survey. (D) None of the above.
2. What is the simple arithmetic mean of 15, 0, 36, 0 and 9 ?
- (A) 20. (B) 15.  
(C) 10. (D) 60.
3. What is the median of the following set of scores 13, 25, 70, 19, 37, 42 ?
- (A) 44.5. (B) 25.  
(C) 31. (D) 37.
4. Which of the following measures of central tendencies need not be possessed by all distributions ?
- (A) Geometric Mean. (B) Harmonic Mean.  
(C) Median. (D) Mode.
5. \_\_\_\_\_ is used when only a rough measure of dispersion is required.
- (A) Mean Deviation. (B) Standard Deviation.  
(C) Quartile Deviation. (D) Range.
6. \_\_\_\_\_ is used to compare the consistency of 2 or more sets of data.
- (A) Co-efficient of Variation. (B) Co-efficient of Correlation  
(C) Co-efficient of Kurtosis. (D) Co-efficient of Skewness.

7. If all values of a sample are same, then its variance is :

- (A) 1. (B) 0.  
(C) 2. (D) Cannot be determined.

8. Which among the following is the possible value of Co-efficient of Correlation between 2 sets of values ?

- (A) - 3. (B) - 1.9.  
(C) - 0.5. (D) 2.

9. Fit a straight line for the following data :

X ... 1 2 3 4 6 8

Y ... 2.4 3 3.6 4 5 6

- (A)  $y = 0.605x + 1.796$ . (B)  $y = 0.605x + 1.976$ .  
(C)  $y = 0.506x + 1.796$ . (D)  $y = 0.506x + 1.976$ .

10. Which of the following can be regression co-efficients  $b_{yx}$  and  $b_{xy}$  ?

- (A) (2, 1). (B) (3, 2/3).  
(C) (6, 1/7). (D) (1/2, 5/2).

11. Which among the following is a sample space obtained while tossing a coin thrice ?

- (A) {(H,T),(T,H),(T,T),(H,H)}.  
(B) {(H,H,H),(H,T,T),(T,T,T)}.  
(C) {(H,H),(T,T)}.  
(D) {(H,H,H),(H,H,T),(H,T,T),(T,H,T),(H,T,H),(T,T,H),(T,H,H),(T,T,T)}

12. In a random experiment of selecting a red bead from a bag with five beads of colours red, white, blue, green and yellow, probability of "getting a red bead" is :

(A)  $1/6$ . (B)  $1/5$ .  
(C)  $3/5$ . (D)  $1/3$ .

13. For any 2 events A and B,  $P(A \cup B) =$

(A)  $P(A) + P(B)$ . (B)  $P(A) + P(B) - P(A \cap B)$ .  
(C)  $P(A) - P(A \cap B)$ . (D)  $P(A) - P(B) - P(A \cap B)$ .

14. What is the probability of the event that an even number appears when tossing a fair dice ?

(A)  $1/2$ . (B)  $1/6$ .  
(C)  $1/4$  (D)  $2/3$ .

15. Given  $P(A) > 0$ ,  $P(B) > 0$ , then  $P(B/A) =$  \_\_\_\_\_.

(A)  $P(A \cap B) / P(A)$ . (B)  $P(B \cap A) / P(B)$ .  
(C)  $P(A \cup B) / P(A)$ . (D)  $P(B \cup A) / P(B)$ .

**FIRST SEMESTER (CBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STA 1C 01—INTRODUCTORY STATISTICS

(2019—2020 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of calculator and Statistical table are permitted.***Part A (Short answer type Questions)***Each question carries 2 marks.**Maximum marks that can be scored from this part is 20.*

1. Expand : (i) CSO ; and (ii) NSSO.
2. Define discrete and continuous data.
3. Define schedule for data collection.
4. Define : (i) Median ; and (ii) Mode.
5. Find the geometric mean of 1, 2, 8 and 16.
6. Define upper and lower outer fences in a box plot.
7. Define co-efficient of quartile deviation.
8. For two variables X and Y, why the regression co-efficients never differ in their signs.
9. One of the regression lines for the variables X and Y is  $2x + 3y - 6 = 0$ . If the mean of X is 3, identify the mean of Y.
10. Define seasonal variation in a time series.
11. Define index numbers.
12. Define Marshall-Edgeworth index number.

**Turn over**

**Part B (Short Essay/Paragraph type Questions)***Each question carries 5 marks.**Maximum marks that can be scored from this part is 30.*

- Write a short note on DES.
- Write any five points to be considered while designing a questionnaire.
- Calculate the harmonic mean of the observations  $1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}$ .
- State principle of least square and explain the fitting of a curve of the form  $y = ab^x$  to the data  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ .
- Derive an expression for the angle between two regression lines for X and Y.
- Define Pearson's co-efficient of correlation. For 12 observations on the variables X and Y, given  $\sum x = 48, \sum y = 60, \sum x^2 = 288, \sum y^2 = 512$ , and  $\sum xy = 384$ . Calculate Pearson's co-efficient of correlation between X and Y.
- Show that Laspayer's and Paache's index numbers are not satisfying factor reversal test of index numbers.

**Part C (Essay type Questions)***Each question carries 10 marks.**Maximum marks that can be scored from this part is 10.*

- Define raw and central moments. Derive an expression for  $r^{\text{th}}$  central moment in terms of raw moments. Explain how skewness and kurtosis are measured using moments.
- Profit after tax (in lakhs) earned by a small scale industry in last 6 years are given as follows :

<i>Year</i>	2015	2016	2017	2018	2019	2020	2021
<i>Profit</i>	6.2	7.4	8.1	7.6	8.4	8.6	8.2

Fit a trend line using least square method and estimate the profit for the year 2025.

Also find three year moving average values using the given data.

**FIRST SEMESTER (CBCSS—UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STA 1B 01—OFFICIAL STATISTICS AND PROBABILITY

(2019—2020 Admissions)

Time : Two Hours and a Half

Maximum : 80 Marks

*Use of calculator and Statistical table are permitted.*

**Part A (Short Answer Type Questions)**

*Each question carries 2 marks.*

*Maximum marks that can be scored from this part is 25.*

1. Expand : (i) MoSPI ; and (ii) NSSTA
2. What is DES ? State any two of its responsibilities.
3. Find the geometric mean of the values 2, 4.5 and 24.
4. The mean mark of 40 students of a class is 36. The mean mark of 24 boys is 32. Find the mean age of the remaining 16 girls.
5. Define mode. How to estimate the mode of data given in a grouped frequency table ?
6. Find the harmonic mean of the numbers 1,  $1/2$  and  $1/3$ .
7. Write the expression for the angle between two regression lines for the variables  $x$  and  $y$ .
8. For three variables  $x$ ,  $y$  and  $z$ , co-efficient of correlations  $r_{xy} = 0.6$ ,  $r_{xz} = 0.8$  and  $r_{yz} = 0.4$ . Calculate the multiple correlation co-efficient of  $x$  on  $y$  and  $z$ .
9. Define : (i) Sample space and (ii) When  $k$  events said to partition of the sample space ?
10. Write the frequency definition of probability. Write one of its merit and demerit.
11. Probabilities of the events  $A$ ,  $B$  and  $A \cap B$  are respectively 0.5, 0.6 and 0.2. Find :  
(i)  $P(A/B^c)$  (ii)  $P(A^c/B^c)$
12. For three independent events  $A$ ,  $B$  and  $C$ , if  $P(A) = 0.5$ ,  $P(B) = 0.3$ ,  $P(C) = 0.2$ , find the probability of the happening of at least one of these events.

**Turn over**

13. Define probability mass function of a discrete random variable and write any two of its properties.
14. Verify whether,  $f(x) = 3e^{-x}$ , for  $0 \leq x < \infty$ , is a probability density function.
15. If the pdf of X is  $f(x) = \frac{1}{4}$ , for  $-2 \leq x < 2$ , find  $P(|X| > 1)$ .

**Part B (Short Essay/Paragraph Type Questions)**

*Each question carries 5 marks.*

*Maximum marks that can be scored from this part is 35.*

16. Prove that the sum of squares of the deviations of the observations is minimum when it is taken about the arithmetic mean of the observations.
17. Define raw and central moments. Prove that the co-efficient of kurtosis based on moments  $\beta_2$  is never less than 1.
18. For two variables X and Y, define correlation and regression analysis. Also write a short note on scatter diagram.
19. Calculate Pearson's co-efficient of correlation between  $x$  and  $y$ , using the following data :
- |       |    |    |    |    |    |    |
|-------|----|----|----|----|----|----|
| $x$ : | 12 | 17 | 18 | 20 | 22 | 28 |
| $y$ : | 36 | 28 | 22 | 21 | 18 | 18 |
20. State and prove Bayes' theorem.
21. Given the distribution function of X as :

$$F(x) = \begin{cases} 0, & x < -1 \\ 0.2, & -1 \leq x < 1 \\ 0.5, & 1 \leq x < 3 \\ 0.8, & 3 \leq x < 5 \\ 1, & x \geq 5 \end{cases}$$

- (i) Sketch the graph of  $F(x)$ ; and (ii) Obtain the probability distribution of X.

22. If  $X$  is a continuous random variable with p.d.f.  $f(x) = \frac{3}{2}x^2$ , for  $-1 < x < 1$ ;  $f(x) = 0$ , elsewhere, find (i)  $P(X^2 < 0.25)$ ; and (ii)  $P(|X| > 0.4)$
23. If  $X$  is a random variable with p.d.f.  $f(x) = 4x^3$ ,  $0 < x < 1$ ; find the p.d.f. of  $Y = 3X$ .

**Part C (Essay Type Questions)**

*Each question carries 10 marks.*

*Maximum marks that can be scored from this part is 20.*

24. For the following data, calculate : (i) Quartile deviation ; (ii) Mean deviation about median ; and (iii) Standard deviation :

Class	0-10	10-20	20-30	30-40	40-50
Frequency :	1	7	22	14	6

25. Derive Spearman's rank correlation co-efficient. Calculate Spearman's co-efficient of rank correlation between  $X$  and  $Y$  based on the following ranks on  $X$  and  $Y$  :

$X$ :	5	3	6	8	4	7	1	2
$Y$	4	1	8	6	7	5	2	3

26. (i) Define conditional probability. For three events,  $A$ ,  $B$  and  $C$ , prove that :

$$P(A \cup B/C) = P(A/C) + P(B/C) - P(A \cap B/C).$$

- (ii) For  $n$  events,  $A_1, A_2, \dots, A_n$ , prove that  $P(A_1 \cap A_2 \cap \dots \cap A_n) \geq \sum_{i=1}^n P(A_i) - (n-1)$ .

27. The probability density function of a random variable  $X$  is given as :

$$f(x) = \begin{cases} \frac{x}{2}, & 0 < x < 1 \\ c, & 1 < x < 2 \\ -\frac{x}{2} + \frac{3}{2}, & 2 < x < 3 \\ 0, & \text{elsewhere} \end{cases}$$

- (i) Find  $c$ ; (ii) Write  $F(x)$ ; and (iii) Find (a)  $P(1 < X < 3)$ ; and (b)  $P(X < 1.5)$  and (c)  $P(X < 2.5 / X > 1.5)$  using  $F(x)$ .

**FIRST SEMESTER (CBCSS—UG) EXAMINATION, NOVEMBER 2021**

## Statistics

## STA 1C 03—DESCRIPTIVE STATISTICS (FOR GEOGRAPHY)

(2021 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of calculator and Statistical table are permitted.***Section A***Answer atleast eight questions.**Each question carries 3 marks.**All questions can be attended.**Overall ceiling 24.*

1. Name any two sources of secondary data.
2. Define population and sample.
3. Define qualitative and quantitative data.
4. Name any four of the main parts of statistical table.
5. Differentiate simple and weighted arithmetic mean.
6. Identify the 5<sup>th</sup> decile and 75<sup>th</sup> percentile of a data with the three quartiles Q1, Q2, and Q3 as 12, 30 and 46 respectively.
7. Define median. Write the expression for finding the median of a grouped data.
8. Define the terms : (i) Dispersion and ; (ii) Quartile deviation.
9. Calculate the mean deviation about mean of 5, 8, 12, 16 and 19.
10. Find the third central moment of a data, if the first three raw moments as 4, 48 and 86.
11. Calculate the second central moment of the values, 6, 4, 8, 10, 7, 5 and 16.
12. Define line of equal distribution in a Lorenz curve.

(8 × 3 = 24 marks)

**Turn over**

**Section B**

*Answer atleast five questions.*

*Each question carries 5 marks.*

*All questions can be attended.*

*Overall ceiling 25.*

13. Mention any five qualities of a good questionnaire.
14. Explain stratified random sampling.
15. Calculate mean deviation about mode of the following data :

<i>Class</i>	0-10	10-20	20-30	30-40	40-50
<i>Frequency</i>	8	14	20	18	10

16. Define partition values. Explain quartiles, deciles, percentiles and their inter relationships.
17. Calculate the first and second raw moments for the following data :
- |                  |       |       |       |       |       |
|------------------|-------|-------|-------|-------|-------|
| <i>Class</i>     | 15-20 | 20-25 | 25-30 | 30-35 | 35-40 |
| <i>Frequency</i> | 4     | 8     | 14    | 7     | 5     |
18. Define central moments. Write the expressions for the second and fourth central moments in terms of raw moments. Find the second raw moment if the mean and variance of a data is 8 and 4.
19. Explain kurtosis and co-efficient of kurtosis based on moments.

(5 × 5 = 25 marks)

**Section C**

*Answer any one question.*

*Each question carries 11 marks.*

20. Define dispersion. Calculate quartile deviation and standard deviation for the following data :

<i>Class</i>	0-10	10-20	20-30	30-40	40-50	50-60
<i>Frequency</i>	6	10	16	13	4	1

21. Explain skewness. Define positively and negatively skewed distributions. Explain : (i) Pearson's co-efficient of skewness ; and (ii) Co-efficient of skewness ; and (ii) Co-efficient of skewness based on moments.

(1 × 11 = 11 marks)

**FIRST SEMESTER (CBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STA 1C 02—DESCRIPTIVE STATISTICS (FOR PSYCHOLOGY)

(2021 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of calculator and Statistical table are permitted.*

**Section A**

*Answer atleast eight questions.*

*Each question carries 3 marks.*

*All questions can be attended.*

*Overall ceiling 24.*

1. Define quantitative and qualitative data.
2. Define relative frequency and frequency density.
3. Define frequency distribution.
4. Differentiate between histogram and bar diagram.
5. What is the Sturges formula while making a frequency table ?
6. Use an appropriate measure to calculate the average speed if the up and down speeds are respectively 30 and 40 km/h.
7. Write any two advantages and disadvantages of median as a measure of average.
8. Calculate the weighted arithmetic mean of the values 20, 25 and 30, if the respective weights are 2, 3 and 4.
9. Define semi inter quartile range.
10. Find the mean deviation about the median of the values 8, 14, 12, 11, 17, 18.
11. Define Bowley's co-efficient of skewness.
12. Define Partition values.

(8 × 3 = 24 marks)

**Turn over**

**Section B**

*Answer atleast five questions.*

*Each question carries 5 marks.*

*All questions can be attended.*

*Overall ceiling 25.*

13. Define primary data. Explain various methods for collecting primary data.
14. Write a short note on frequency curve and frequency polygon.
15. Define ogives. Construct a greater than ogive for the following data :

<i>Class</i>	0-10	10-20	20-30	30-40	40-50
--------------	------	-------	-------	-------	-------

<i>Frequency</i> :	6	10	14	8	2
--------------------	---	----	----	---	---

16. The average accidents reported in a city for the first 10 days of April is 4 per day and for the remaining days ; it is noted as 2.5 per day. Calculate the average daily accidents for the entire month.
17. Find the quartile deviation for the following data :

<i>Class</i>	0-10	10-20	20-30	30-40	40-50
--------------	------	-------	-------	-------	-------

<i>Frequency</i> :	6	10	20	12	8
--------------------	---	----	----	----	---

18. Two series A and B of data are given. Which series shows more variability ?
19. Explain various type of skewness and Pearson's measure of skewness.

(5 × 5 = 25 marks)

**Section C**

*Answer any one question.*

*Each question carries 11 marks.*

20. Define dispersion. Calculate the : (i) Mean deviation about mean ; and (ii) Standard Deviation for the following frequency distribution :

<i>Class</i> :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
----------------	------	-------	-------	-------	-------	-------	-------	-------

<i>Frequency</i> :	7	13	18	26	20	10	4	2
--------------------	---	----	----	----	----	----	---	---

21. Explain kurtosis. Calculate percentile measure of kurtosis for the following data :

<i>Class</i>	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
--------------	------	-------	-------	-------	-------	-------	-------	-------

<i>Frequency</i>	7	13	18	26	20	10	4	2
------------------	---	----	----	----	----	----	---	---

(1 × 11 = 11 marks)

**FIRST SEMESTER (CBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STA 1C 01—INTRODUCTORY STATISTICS

(2021 Admissions)

(Multiple Choice Questions for SDE Candidates)

**Time : 15 Minutes****Total No. of Questions : 15****Maximum : 15 Marks****INSTRUCTIONS TO THE CANDIDATE**

1. This Question Paper carries Multiple Choice Questions from 1 to 15.
2. The candidate should check that the question paper supplied to him/her contains all the 15 questions in serial order.
3. Each question is provided with choices (A), (B), (C) and (D) having one correct answer. Choose the correct answer and enter it in the main answer-book.
4. The MCQ question paper will be supplied after the completion of the descriptive examination.

## STA 1C 01—INTRODUCTORY STATISTICS

(Multiple Choice Questions for SDE Candidates)

1. A study based on complete enumeration of data is known as :
- (A) Sample survey. (B) Pilot survey.  
(C) Census survey. (D) None of the above.
2. Which among the following is a measure of positional average ?
- (A) Arithmetic mean. (B) Median.  
(C) Harmonic mean. (D) Geometric mean.
3. What is the median of the following set of scores 13, 25, 70, 19, 37, 42 ?
- (A) 44.5. (B) 25.  
(C) 31. (D) 37.
4. The value in a distribution which occurs most frequently is \_\_\_\_\_.
- (A) Arithmetic Mean. (B) Geometric Mean.  
(C) Median. (D) Mode.
5. What is the geometric mean of 2, 4 and 8 ?
- (A) 4. (B) 6.  
(C) 5. (D) 7.
6. Mean marks of 150 students were found to be 50. Later it was found that a score of 87 was misread as 78. What is the correct mean ?
- (A) 49.9. (B) 59.  
(C) 50.09. (D) None of these.

7. Which among the following is equal to the measure at the  $(N + 1)/2^{\text{th}}$  position of an ordered data ?
- (A) Median. (B) 2nd Quartile.  
(C) Both (A) and (B). (D) Neither (A) nor (B).
8. Calculate the standard deviation for the following data :
- 5, 8, 7, 11, 9, 10, 8, 2, 4, 6 :
- (A) 2. (B)  $\sqrt{5}$ .  
(C)  $\sqrt{6}$ . (D)  $\sqrt{7}$ .
9. \_\_\_\_\_ analysis deals with the intensity of relationship among variables.
- (A) Regression. (B) Correlation.  
(C) Both (A) and (B). (D) None of these.
10. Which among the following is not possible value of co-efficient of correlation between 2 sets of values ?
- (A) - 3. (B) -1.9.  
(C) - 0.5. (D) 2.
11. Which of the following can be regression co-efficients  $b_{yx}$  and  $b_{xy}$  :
- (A) (2, 1). (B) (3, 2/3).  
(C) (6, 1/7). (D) (1/2, 5/2).

12. Which of the given sets are not an event of the sample space  $S = \{a, 1, 2, 3, z\}$  :
- (A)  $\{a, z\}$ . (B)  $\{1, 2, 3\}$ .  
(C)  $\{\}$ . (D)  $\{a, b\}$ .
13. In a random experiment of rolling a die and observing the number shown up, let  $A$  be the event "odd number showing up". Then  $A^c =$  :
- (A)  $\{1, 2, 3, 6\}$ . (B)  $\{1\}$ .  
(C)  $\{1, 3, 5\}$ . (D)  $\{2, 6\}$ .
14. Phenomenon of Statistical Regularity is observed when :
- (A) Number of trails of a random experiment increases.  
(B) Number of trails of a random experiment is kept minimal.  
(C) Relative frequencies approach divergent values.  
(D) None of these.
15. For any 2 events  $A$  and  $B$ ,  $P(A \cup B) =$
- (A)  $P(A) + P(B)$ . (B)  $P(A) + P(B) - P(A \cap B)$ .  
(C)  $P(A) - P(A \cap B)$ . (D)  $P(A) - P(B) - P(A \cap B)$ .

**FIRST SEMESTER (CBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STA 1C 01—INTRODUCTORY STATISTICS

(2021 Admissions)

Time : Two Hours

Maximum : 60 Marks

*Use of Calculator and Statistical tables are permitted.*

**Section A**

*Answer atleast eight questions.*

*Each question carries 3 marks.*

*All questions can be attended.*

*Overall ceiling 24.*

1. Expand CSO and mention any two of its responsibilities.
2. Distinguish between questionnaire and schedule.
3. Name any four different types of bar diagrams.
4. If the variance of  $x_1, x_2, \dots, x_n$  is  $k$ , identify the variances of the sets :  
(i)  $x_1 - 5, x_2 - 5, \dots, x_n - 5$ ; (ii)  $5x_1, 5x_2, \dots, 5x_n$ .
5. Find the mean deviation about median of the observations, 4, 7, 15, 12, 10 and 18.
6. Define H-spread.
7. Second, third and fourth central moments of a data are 5.2, 2 and 30 respectively. Obtain the co-efficients of skewness and kurtosis.
8. What are the regression co-efficients and state their relation between Pearson's co-efficient of correlation for two variables X and Y ?
9. Comment on the co-efficient of correlation between two variables X and Y, if the angle between the regression lines : (i)  $0^\circ$ ; and (ii)  $90^\circ$ .
10. Define cyclical and irregular variation in a time series data.

**Turn over**

11. Define "base year" and "current year" while constructing index numbers.
12. Define Laaspayer's and Paasche's price index numbers.

(8 × 3 = 24 marks)

**Section B***Answer atleast five questions.**Each question carries 5 marks.**All questions can be attended.**Overall ceiling 25.*

13. Write a short note on the statistical system in India.
14. Differentiate between primary and secondary data. Explain various methods for collecting primary data.
15. Explain kurtosis and its measure based on moments.
16. State the principle of least squares for curve fitting. Explain the method of fitting a curve  $y = ae^{bx}$  using the observations  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  and  $x$  and  $y$ .
17. Obtain the regression line  $x$  on  $y$  and regression line  $y$  on  $x$  using the following data on  $x$  and  $y$  :
 

$x$	8	12	15	17	20	23	25
$y$	10	15	16	20	25	30	35
18. Show that Pearson's co-efficient of correlation is invariant under linear transformation.
19. Explain the method of semi-average for finding the secular trend in a time series data.

(5 × 5 = 25 marks)

**Section C***Answer any one question.**The question carries 11 marks.*

20. (i) Define Dispersion. Calculate the mean deviation about median for the following data :

<i>Class</i>	5-15	15-25	25-35	35-45	45-55	55-65	65-75
<i>Frequency</i>	4	11	19	30	10	4	2

- (ii) Define raw and central moments. Express  $r^{\text{th}}$  central moment in terms of raw moments and hence obtain the expression for fourth central moment.
21. (i) Define Index Numbers. Explain various types of index numbers.
- (ii) Calculate Fisher's index number for the following data :

Items	Quantity		Price	
	2016	2021	2016	2021
A	12	18	18	24
B	14	17	18	22
C	11	12	15	14
D	19	24	26	26
E	8	10	12	17

(1 × 11 = 11 marks)

**FIRST SEMESTER (CBCSS-UG) DEGREE EXAMINATION, NOVEMBER 2021**

Statistics

STA 1B 01—OFFICIAL STATISTICS AND PROBABILITY

(2021 Admissions)

Time : Two Hours and a Half

Maximum : 80 Marks

**Section A (Short Answer Questions)***Answer at least ten questions.**Each question carries 3 marks.**All questions can be attended.**Overall Ceiling 30.*

1. Expand (i) MoSPI ; (ii) CSO.
2. State any two functions of DES.
3. Define central tendency.
4. Find the geometric mean of 1, 2 and 32.
5. Find the variance of a set of values with mean 10 and coefficient of variation 50%.
6. What are the interrelations between quartiles and percentiles ?
7. Write any two properties of regression of coefficients for the variables X and Y.
8. Define partial correlation when three variables X, Y and Z are considered.
9. Define mutually exclusive events. Site an example.
10. Using the three events A, B and C, express the events (i) at least one of A, B and C ; (ii) none of A, B and C, in set notation.
11. Given  $P(A) = 0.8, P(B) = 0.5$  and  $B \subset A$ . Find  $P(A \cap B^c)$ .
12. Given  $P(A) = 0.4, P(A \cup B) = 0.7$ . Find  $P(B)$ , if  $P(B/A) = 0.6$ .
13. Define cumulative distribution function of the r.v., X and write any two of its properties.
14. Find  $P(0.3 < X < 0.8)$ , if the p.d.f. of X is,  $f(x) = 2x$ , for  $0 < x < 1$ ;  $f(x) = 0$ , elsewhere.
15. Write the p.m.f.  $f(x)$  of the r.v., X, if the distribution function of X is :  
 $F(x) = 0$ , for  $x < 1$ ;  $F(x) = 0.3$ , for  $1 < x < 2$ ;  $F(x) = 1$ , for  $x > 2$ .

(10 × 3 = 30 marks)

**Turn over**

## Section B (Paragraph Type Questions)

*Answer at least five questions.*

*Each question carries 6 marks.*

*All questions can be attended.*

*Overall Ceiling 30.*

16. Show that the sum of squares of the deviation of a set of observations is minimum, when it is taken about the arithmetic mean of the set.
17. Calculate the 4<sup>th</sup> decile and 70<sup>th</sup> percentile for the following data :

Class	0-10	10-20	20-30	30-40	40-50
Frequency	15	26	38	13	8

18. Explain the least square method of fitting of the parabola  $y = ax^2 + bx + c$ , for a given set of observations  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ .
19. For the two regression lines  $3y - 4x = 14$  and  $16x = 4y + 8$ . Identify the regression coefficients, find the means of  $x$  and  $y$  and the coefficient of correlation.
20. State the frequency definition of probability and using this establish  $P(A^c) = 1 - P(A)$ .
21. State Bayes' theorem. Out of the total e-mails received in an account, 30% are spam. 80% of these spam messages contains the word 'offer'. Also 10% of desired emails contain the word 'offer'. What is the probability that a new email containing the word 'offer' is a spam ?
22. Define probability density function of a random variable  $X$ . State any two of its properties. Also examine whether the following function is a probability density function

$$f(x) = \frac{2x^2}{3}, 0 < x < 2; f(x) = 0, \text{ elsewhere.}$$

23. If  $X$  is a random variable with p.d.f.  $f(x) = e^{-x}, 0 < x < \infty$ ; find the p.d.f. of  $Y = 1 - e^{-x}$ .

(5 × 6 = 30 marks)

## Section C (Essay type Questions)

Answer any two questions.  
Each question carries 10 marks.

24. (i) Define coefficient of variation. Calculate the coefficient of variation for the following frequency distribution :

Class	0-10	10-20	20-30	30-40	40-50
Frequency	5	7	8	6	4

- (ii) Define kurtosis and explain any two of its measures in detail.
25. (i) Define rank correlation coefficient. Derive Spearman's rank correlation coefficient.
- (ii) Following are the scores for 10 students based on their performance in co-curricular activities (A) and academic activities (B) Calculate Spearman's rank correlation coefficient between A and B :

Students	1	2	3	4	5	6	7	8	9	10
Scores in A	39	45	35	29	28	40	42	43	30	25
Scores in B	43	38	40	34	36	44	35	41	37	39

26. (i) State mathematical definition of the probability of an event. Content of box of 10 identical balls is 3 red and 7 black balls. Balls are randomly selected one by one and noted their colour. What is the probability that the fifth ball selected is the last red ball ?
- (ii) Define independence of two events. If A and B are two independent events, show that their complement events are also independent .
27. The probability mass function of a random variable X is given as,

$$f(x) = \begin{cases} \frac{2}{9}, & \text{when } x = 0 \\ k, & \text{when } x = 2 \\ \frac{4}{9}, & \text{when } x = 4 \\ 0, & \text{otherwise} \end{cases}$$

- (i) Find  $k$ .
- (ii) Determine the Distribution function  $F(x)$  and sketch the graph of  $F(x)$ .
- (iii) Using  $F(x)$ , identify (i)  $P(X > 2)$ ; (ii)  $P(1.8 < X < 2)$ ; (iii)  $P(2 < X < 2.5)$ .

(2 × 10 = 20 marks)

FIRST SEMESTER (CUCBCSS—UG) DEGREE EXAMINATION  
NOVEMBER 2021

Statistics

SG 1C 01—STATISTICAL MECHANICS

(2016—2018 Admissions)

Time : Three Hours

Maximum : 80 Marks

*Use of calculator is permitted.*

**Section A (One Word Questions)**

*Answer all questions.*

*Each question carries 1 mark.*

1. Complete enumeration of a population is known as \_\_\_\_\_.
2. The first time data collected by the investigator directly is called \_\_\_\_\_.
3. The value obtained by dividing class frequency by class width is known \_\_\_\_\_.
4. \_\_\_\_\_ is a positional average.
5. Arithmetic mean of a set of 9 observations is 12. When an observation 7 is added to this group the arithmetic mean becomes \_\_\_\_\_.
6. X- Coordinate of the meeting point of less than and greater than frequency curves is \_\_\_\_\_ of the data.
7. Difference between highest and least observations in a set is known as \_\_\_\_\_.
8. The standard deviation of a set of values where all the values are same is \_\_\_\_\_.
9. 40% of the observations of a set are greater than 72 and 60 % of the observations are greater than 58. Then the 6<sup>th</sup> decile of the set of observations is \_\_\_\_\_.
10. Mode of a symmetric observation with mean 14 is \_\_\_\_\_.

(10 × 1 = 10 marks)

**Section B (One Sentence Questions)**

*Answer all questions.*

*Each question carries 2 marks.*

11. Differentiate qualitative and quantitative data.
12. Define relative frequency of a class.

13. Define cumulative frequency distribution.
14. Two sets of observations of sizes 20 and 30 are having means 10 and 15 respectively. If these two sets are combined what is the mean of the new set ?
15. Define Mode.
16. Define Dispersion.
17. Find the quartile co-efficient of dispersion of a set where the first and last quartiles are 2 and 18 respectively.

(7 × 2 = 14 marks)

**Section C (Paragraph Questions)**

*Answer any three questions.  
Each question carries 4 marks.*

18. Illustrate the difference between frequency polygon and frequency curve for the following data :

Class	0-10	10-20	20-30	30-40	40-50
Frequency	3	5	9	6	2

19. Define secondary data. Mention any *two* sources for secondary data.
20. Calculate the median for the following data :

Class	0-5	5-10	10-15	15-20	20-25
Frequency	7	10	16	11	6

21. Obtain mean deviation about mean of 10, 12, 25, 14, 18, 24 and 23.
22. Explain Lorenz curve as a graphical measure of dispersion.

(3 × 4 = 12 marks)

**Section D (Short Essay Questions)**

*Answer any four questions.  
Each question carries 6 marks.*

23. Explain any four methods for collecting primary data.
24. Define Sampling. Mention any *five* advantages of sampling over census.

25. Explain the method of drawing a pie diagram to the marks scored in four papers by a student given below :

Subject	Mathematics	Physics	Chemistry	Biology	Total
Marks	64	82	54	90	290

26. Explain the desirable properties of a good average.  
 27. Obtain the standard deviation of first  $n$  natural numbers.  
 28. Define Skewness. Explain various types of skewness and Bowley's co-efficient of skewness.

(4 × 6 = 24 marks)

### Section E (Essay Questions)

*Answer any two questions.*

*Each question carries 10 marks.*

29. Define bar diagram. What are the different types of bar diagram ? Illustrate with suitable examples.  
 30. Calculate the Geometric mean and Harmonic mean for the following data :

Class	0-10	10-20	20-30	30-40	40-50
Frequency	4	8	10	6	7

31. (i) Define co-efficient of variation.  
 (ii) Calculate the co-efficient of variation for the following data :

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequency	1	2	9	8	5	4	1

32. (i) Define Kurtosis.  
 (ii) Calculate the co-efficient of kurtosis based on partition values for the following data :

Class	1-5	6-10	11-15	16-20	21-25	26-30	31-35
Frequency	3	4	68	30	10	6	2

(2 × 10 = 20 marks)

**FIRST SEMESTER (CUCBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STS 1C 01—BASIC STATISTICS AND PROBABILITY

(2016—2018 Admissions)

(Multiple Choice Questions for SDE Candidates)

**Time : 15 Minutes**

**Total No. of Questions : 20**

**Maximum : 20 Marks**

**INSTRUCTIONS TO THE CANDIDATE**

1. This Question Paper carries Multiple Choice Questions from 1 to 20.
2. The candidate should check that the question paper supplied to him/her contains all the 20 questions in serial order.
3. Each question is provided with choices (A), (B), (C) and (D) having one correct answer. Choose the correct answer and enter it in the main answer-book.
4. The MCQ question paper will be supplied after the completion of the descriptive examination.

## STS 1C 01—BASIC STATISTICS AND PROBABILITY

(Multiple Choice Questions for SDE Candidates)

1. The totality of all objects under a study is called \_\_\_\_\_.
- (A) Sample. (B) Group.  
(C) Population. (D) Specimen.
2. A study based on complete enumeration of data is known as :
- (A) Sample survey. (B) Pilot survey.  
(C) Census survey. (D) None of the above.
3. What is the simple arithmetic mean of 15, 0, 36, 0 and 9 ?
- (A) 20. (B) 15.  
(C) 10. (D) 60.
4. What is the median of the following set of scores 13, 25, 70, 19, 37, 42 ?
- (A) 44.5. (B) 25.  
(C) 31. (D) 37.
5. Median =  $1 + c(N/2 - m)/f$  where 'm' is :
- (A) Cumulative frequency of the median class.  
(B) Frequency of the median class.  
(C) Cumulative frequency of the class preceding the median class.  
(D) Frequency of the class preceding the median class.
6. That value in a distribution which occurs most frequently is \_\_\_\_\_.
- (A) Arithmetic Mean. (B) Geometric Mean.  
(C) Median. (D) Mode.
7. What is the geometric mean of 2, 4 and 8 ?
- (A) 4. (B) 6.  
(C) 5. (D) 7.

8. Population of a state increased by 9 %, 16 % respectively from year 2012 to 2013 What is the average increase over these years ?
- (A) 12.5 %. (B) 9 %.  
(C) 12 %. (D) 14 %.
9. Which among the following is not a commonly used measure of dispersion ?
- (A) Range. (B) Median.  
(C) Standard Deviation. (D) Mean Deviation.
10. \_\_\_\_\_ is used to compare the consistency of 2 or more sets of data.
- (A) Co-efficient of Variation.  
(B) Co-efficient of Correlation.  
(C) Co-efficient of Kurtosis.  
(D) Co-efficient of Skewness.
11. \_\_\_\_\_ is the best measure of dispersion.
- (A) Standard Deviation. (B) Quartile Deviation.  
(C) Mean Deviation. (D) Range.
12. If the minimum value in a set is 12 and its range is 8, what is its maximum value ?
- (A) 4. (B) 20.  
(C) 12. (D) 8
13. Rank Correlation was obtained by \_\_\_\_\_ to measure degree of relationship between qualitative characteristics.
- (A) Karl Pearson. (B) Galton.  
(C) Spearman. (D) None of them.
14. Which among the following is the possible value of Co-efficient of Correlation between 2 sets of values ?
- (A) - 3. (B) - 1.9.  
(C) - 0.5. (D) 2.

15. Which of the following is not a characteristic of a random experiment ?
- (A) Number of outcomes is 2 or more.
  - (B) All outcomes are not known in advance.
  - (C) Outcome obtained in a particular trial is not known in advance.
  - (D) Experiment can be repeated under identical conditions.
16. Classical definition of probability gives that if the elementary events of a random experiment are mutually exclusive, exhaustive and equally likely, then :
- (A)  $P(A) = \text{Number of outcomes of } A / \text{Total number of outcomes.}$
  - (B)  $P(A) = \text{Number of elements in } A.$
  - (C)  $P(A) = \text{Number of elements in } S / \text{Number of outcomes of } A.$
  - (D) None of the above.
17. Phenomenon of Statistical Regularity is observed when :
- (A) Number of trials of a random experiment increases.
  - (B) Number of trials of a random experiment is kept minimal.
  - (C) Relative frequencies approach divergent values.
  - (D) None of these.
18. For any 2 events A and B,  $P(A \cup B) =$
- (A)  $P(A) + P(B).$
  - (B)  $P(A) + P(B) - P(A \cap B).$
  - (C)  $P(A) - P(A \cap B).$
  - (D)  $P(A) - P(B) - P(A \cap B).$
19. Which of these is the same as  $(A \cap B^c) \cup (A \cap B)$ ?
- (A)  $A \setminus B.$
  - (B)  $A \cup B.$
  - (C)  $\bar{A}.$
  - (D)  $A \cap B.$
20. What is the probability that there should be 53 Sundays in a non-leap year ?
- (A)  $2/7.$
  - (B)  $1/7.$
  - (C)  $53/365.$
  - (D)  $312/365.$

**FIRST SEMESTER (CUCBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021**

Statistics

STS 1C 01—BASIC STATISTICS AND PROBABILITY

(2016—2018 Admissions)

Time : Three Hours

Maximum : 80 Marks

**Section A**

*Answer all questions in one word.*

*Each question carries 1 mark.*

Fill up the blanks :

1. A part of the population is known as \_\_\_\_\_.
2. The sum of the deviations from mean is always \_\_\_\_\_.
3. If  $r = 0$ , the two lines of regression are at an angle of \_\_\_\_\_.
4. If  $B \subset A$ , then probability of  $(A \cap \bar{B})$  is \_\_\_\_\_.
5. \_\_\_\_\_ can be calculated from a frequency distribution with open end classes.

Write true or false :

6. If mean and standard deviation of a distribution are 20 and 4 respectively, co-efficient of variation is 15%.
7. The geometric mean of a set of values lies between arithmetic mean and harmonic mean.
8. The value of multiple correlation co-efficient lies in between  $-1$  and  $+1$ .
9. If A and B are two events which have no events in common, then the events A and B are independent.
10. For a continuous random variable X,  $P(a < X < b) = F(b) - F(a)$ , Where  $F(\cdot)$  is the distribution function of X.

(10 × 1 = 10 marks)

**Turn over**

## Section B

Answer all questions in one sentence each.

Each one carries 2 marks.

11. Define mode.
12. What do you mean by quartiles ?
13. If  $r_{12} = 0.4$ ,  $r_{23} = 0.5$  and  $r_{13} = 0.6$ , calculate multiple correlation co-efficient  $R_{1,23}$ .
14. Suppose A and B are any two events and  $P(A) = p_1$ ,  $P(B) = p_2$  and  $P(A \cap B) = p_3$ . Find  $P(\bar{A} \cup B)$ .
15. State the axiomatic definition of probability.
16. Let X be a random variable with pdf  $f(x) = \begin{cases} ke^{-2x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$ , find the value of k.
17. Explain the term statistical regularity.

(7 × 2 = 14 marks)

## Section C

Answer any three questions.

Each question carries 4 marks.

18. The arithmetic mean of two observations is 127.5 and their GM is 60. Find : i) Harmonic mean and ; ii) The two observations.
19. State and prove addition theorem of probability.
20. Prove with an example that pairwise independence does not imply mutual independence.
21. Three newspapers A B and C are published in a city. It is estimated from a survey that of the adult population : 20% read A, 16% read B, 14% read C, 8% read both A and B, 5% read both A and C, 4% read both B and C, 2% read all three. Find what percentage read at least one of the papers ?
22. If X is a random variable with pmf  $p(x) = pq^x$ ,  $x = 0, 1, 2, \dots$ ,  $0 < p < 1$ ,  $p + q = 1$ , find the distribution of  $Y = 2X$ .

(3 × 4 = 12 marks)

## Section D

Answer any four questions.

Each one carries 6 marks.

23. Compute quartile deviation and co-efficient of quartile deviation for the following data :

Variable	10	20	30	40	50	60
Frequency	4	7	15	8	7	2

24. Using the principle of least squares explain the procedure for fitting the curve  $Y = ab^x$ .

25. Calculate Karl Pearson's co-efficient of correlation for the following data :

X	6	8	12	15	18	20	24	38	31
Y	10	12	15	15	18	25	22	26	28

26. If  $f(x) = \frac{x}{15}$ ,  $x = 1, 2, 3, 4, 5$ , find : i)  $P(x = 1 \text{ or } x = 2)$  and ii)  $P(1/2 < X < 5/2 | X > 1)$ .

27. If A and B are independent events, show that : i)  $\bar{A}$  and  $\bar{B}$  are independent ; ii)  $\bar{A}$  and B are independent iii) A and  $\bar{B}$  are independent.

28. Let X be a continuous random variable with pdf  $f(x) = e^{-x}$ ,  $0 < x < \infty$ . Find the probability density function of : i)  $Y = 2X + 5$  ; and ii)  $Y = X^3$ .

(4 × 6 = 24 marks)

## Section E

Answer any two questions.

Each question carries 10 marks.

29. i) Define standard deviation.  
ii) Calculate the co-efficient of variation of the following two series and find which series is more consistent.

Variable	10-20	20-30	30-40	40-50	50-60	60-70
Series A	10	18	32	40	22	18
Series B	18	22	40	32	18	10

Turn over

30. Obtain the equation of the line of regression of Y on X and X on Y. Show that the angle  $\theta$  between the two lines of regression is given by  $\tan \theta = \left( \frac{r^2 - 1}{r} \right) \frac{\sigma_x \sigma_y}{\sigma_x^2 \sigma_y^2}$  where  $r$  is the correlation co-efficient between X and Y.

31. i) State and prove Bayes theorem.  
 ii) The contents of urns I, II and III are as follows :

Urn I            1 white, 2 black and 3 red balls

Urn II           2 white, 1 black and 1 red balls

Urn III ...    4 white, 5 black and 3 red balls

One urn is chosen at random and two balls drawn from it. They happen to be white and red. What is the probability that they come from urns I, II or III ?

32. Let X be a continuous random variable with pdf  $f(x) = 6x(1-x)$ ,  $0 \leq x \leq 1$ .

- i) Check that the above is a pdf.  
 ii) Obtain the distribution function of X.  
 iii) Compute  $P\left(X \leq \frac{1}{2} \mid \frac{1}{3} \leq X \leq \frac{2}{3}\right)$ .  
 iv) Determine the number  $k$  such that  $P(X < k) = P(X > k)$ .

(2 × 10 = 20 marks)

FIRST SEMESTER (CUCBCSS-UG) DEGREE EXAMINATION  
NOVEMBER 2021

Statistics

STS 1B 01—BASIC STATISTICS AND PROBABILITY

(2016—2018 Admissions)

Time : Three Hours

Maximum : 80 Marks

*Use of Calculator is permitted.***Section A (One Word Questions)***Answer all questions.**Each question carries 1 mark.*

1. The mean of a set of 25 observations when one of the observations 65 was misread as 40 is 40. The corrected mean is \_\_\_\_\_.
2. Sum of the absolute values of the deviations of the observations is minimum when it is taken about the \_\_\_\_\_ of the observations.
3. A constant 5 is added to all observations of a set with standard deviation 6. The standard deviation of the new set is \_\_\_\_\_.
4. For an event E,  $P(E) + P(E^C) =$  \_\_\_\_\_.
5. The set of all possible outcomes of a random experiment is known as \_\_\_\_\_.
6.  $P(A \cup B)$  for two independent events with  $P(A) = 0.6$  and  $P(B) = 0.4$  is \_\_\_\_\_.
7. A and B are events of a sample space S.  $P(A) = 0.5$ ,  $P(B) = 0.3$ ,  $P(A \cap B) = 0.2$ , then  $P(\text{At least one of A and B occurs}) =$  \_\_\_\_\_.
8. Given the p. m. f.,  $f(x) = k$ , for  $x = 0, 1, 2, 3, 4$  and  $f(x) = 0$ , elsewhere. Then  $k =$  \_\_\_\_\_.
9. Given the p.d.f.  $f(x) = \frac{1}{5}$ ,  $3 < x < 8$  and  $f(x) = 0$ , otherwise, then  $P(1 < X < 5) =$  \_\_\_\_\_.
10. The regression coefficient  $x$  on  $y$  when the regression line  $x$  on  $y$  is  $10x + 6y - 12 = 0$  is \_\_\_\_\_.

(10 × 1 = 10 marks)

**Turn over**

**Section B (One Sentence Questions)***Answer all questions.**Each question carries 2 marks.*

11. Obtain the harmonic mean of 3, 4 and 5.
12. Define deciles and percentiles.
13. Define statistical regularity.
14. Prove that  $P(A/B) + P(A^c/B) = 1$ .
15. Verify whether  $p(x) = \frac{x^2}{15}$ ,  $x = -1, 0, 1, 2, 3$  is a probability mass function.
16. Differentiate between discrete and continuous random variables.
17. Write any two properties of regression co-efficients.

**(7 × 2 = 14 marks)****Section C (Paragraph Questions)***Answer any three questions.**Each question carries 4 marks.*

18. State and prove any two properties of arithmetic mean.
19. Calculate the quartile deviation for the following data :
 

Class	0-10	10-20	20-30	30-40	40-50
Frequency	16	22	35	17	10
20. Define mean deviation. Calculate mean deviation about mean of the following data :

X :	2	4	6	8	10
f	1	4	6	4	1

21. Obtain the distribution function of X with p.m.f.  $p(x) = \frac{x+2}{9}$ ,  $x = 0, 1, 2$  and  $p(x) = 0$ , elsewhere.
22. Write a short note on partial and multiple correlations when three variables are considered.

**(3 × 4 = 12 marks)**

30. State and prove Bayes' theorem. In answering a question on a multiple choice test with four choices, an examinee either knows the answer with probability 0.4 or he guesses with probability 0.6. Let the probability of answering the question correctly is 1 for an examinee who knows the answer and  $\frac{1}{4}$  who guesses. Suppose an examinee answer a question correctly. What is the probability that he really knows the answer ?
31. Given the p.d.f. of a continuous r.v.  $X$  as,  $f(x) = k(x+1)$ ,  $-1 < x < 2$ ;  $f(x) = 0$ , otherwise. Obtain : (i) The value of  $k$  ; (ii) The distribution function  $F(x)$  ; (iii)  $P(-1 < X < 1)$  ; and (iv)  $P(X < 1/X > 0)$  ; and ; (v)  $P(X > 0/X < 1)$ .
32. Explain curve fitting. Derive normal equations for fitting a straight line of the form  $y = ax + b$  using the principle of least squares for the given set of observations  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . Making use of these normal equations, explain fitting of a curve of the form  $y = ax^b$  for the given data.

(2 × 10 = 20 marks)

**Section D (Short Essay Questions)***Answer any four questions.**Each question carries 6 marks.*

23. If A and B are two independent events prove that : (i) A and  $B^c$ ; (ii)  $A^c$  and B are also independent.
24. For two events,  $P(A) = p_1$ ,  $P(B) = p_2$  and  $P(A \cap B) = p_3$ , where  $p_i$ 's are  $\geq 0$ . Find the expressions for :
- (i)  $P(A \cup B)^c$ ; (ii)  $P(A^c \cup B)$ ; and (iii)  $P(A^c / B^c)$ .
25. Write the classical definition of probability. In a class there are 3 boys and 2 girls. 3 students are selected at random from the class. Find the probability that the selected group contains 2 boys and 1 girl or 1 boy and 2 girls.
26. If the random variable X takes the values 0, 1, 2, 3 and 4 such that,
- $$3P(X = 0) = \frac{3}{2}P(X = 1) = P(X = 2) = P(X = 3) = 3P(X = 4).$$
- Find the probability distribution of X.
- Also write the probability distribution of  $Y = X - 4$ .
27. State and prove the multiplication theorem on probability for : (i) any three events A, B and C ; and (ii) The independent events A, B and C.
28. Derive Spearman's rank correlation co-efficient.

**(4 × 6 = 24 marks)****Section E (Essay Questions)***Answer any two questions.**Each question carries 10 marks.*

29. Define coefficient of variation.

From the following two series of data, find which series is more consistent :

Class	5-10	10-15	15-20	20-25	25-30	30-35
Series A :	2	9	29	54	11	5
Series B :	9	11	18	32	27	13

**Turn over**