1 SEM MCOM (CBCS) STS 200 (N)

2019

(December)

COMMERCE

Paper : 10200

(Statistics)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer ***any five*** of the following : 2×5=10

   (a) Define Aximatic Definition of Probability.

   (b) What is null hypothesis and alternative hypothesis ?

   (c) Differentiate with example between Simple Linear Regression Model and Multiple Linear Regression Model.

(d) Under what conditions Non-parametric methods are more suitable to use than Parametric methods.

(e) Under what conditions Poisson distribution can be considered as a limiting case of binomial distribution.

(f) Name *at least four* Probability Sampling Method.

2. Answer ***any four*** questions : 4×4=16

(a) Discuss the basic principles of sample survey.

(b) Describe the errors in Hypothesis testing.

(c) Discuss Simple Random Sampling and Stratified Random Sampling.

(d) Define one-tailed test and two-tailed test.

(e) Ten coins are thrown simultaneously. Find the probability of getting at least seven heads using binomial probability law.

(f) Explain the advantages of sampling over complete endumeration.

3. (a) A problem of Statistics is given to three students Amar, Akbar and Anthony whose chances of solving it are $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$ respectively. What is the probability that (*i*) the problem will not be solved and (*ii*) the problem will be solved? 4

Or

(b) Out of 100 students in a boys hostel 80 take tea, 40 take coffee and 25 take both. Find the probability that a student takes (*i*) either tea *or* coffee (*ii*) neither tea *nor* coffee. 4

4. (a) Explain Multiple linear regression model with its assumptions. 5

Or

(b) To develop better work schedules management would like to develop an estimated regression equation that will help predict total daily travel time for its drivers. A random sample of 10 days of operation was taken, the data obtained are presented in the following table. Fit a regression line for the following data : 5

Day		1	2	3	4	5	6	7	8	9	10
Miles travelled	:	100	50	100	50	80	75	80	90	90	100
Travel time	:	9.3	4.8	8.9	4.2	6.8	6.6	5.9	7.6	6.1	5.8

Contd.

5. *(a)* In a factory which manufactures bolts, machine A, B and C manufacture respectively 25%, 35% and 40% of bolts. Of their outputs 5, 4 and 2 per cent are defective bolts. A bolt is drawn at random and found to be defective. What are the probabilities that it was manufactured by the machines A, B and C? 5

Or

(b) The odds that a book will be favourably reviewed by 3 independent critics are 5 : 2, 4 : 3 and 3 : 4 respectively. What is the probability that out of the three reviews a majority will be preferable? 5

6. *(a)* Two salesman A and B are working in a certain district. From a sample survey conducted by the Head office, the following results were obtained. State whether there is any significant difference in the average sales between the two salesman : 5

	A	B
No. of sales	20	18
Average sales (in Rs.)	170	205
Standard Deviation (in Rs.)	20	25

(Table value of t for 36 d.f. at 5% level of significance for a two-tailed test is 2.45)

(b) Suppose it is desired to determine whether 10 students of class X can perform equally in two different tests. The test scores are as follows : 5

Test 1	:	90	90	80	90	92	88	90	63	70	54
Test 2	:	84	84	82	94	90	85	89	62	65	52

(The tabulated value of t for 5% level of significance for α d.f. is 2.262)

7. *(a)* Two random samples drawn from two normal populations are : 5

Sample I : 20, 16, 26, 27, 23, 32, 18, 24, 25, 19

Sample II : 27, 33, 42, 35, 32, 34, 38, 28, 41, 43, 30, 37

Test whether the population have same variances.

(Table value of F for (11, 9) d.f. at 5% level of significance is 3.11)

Or

(b) What do you understand by non-parametric methods of testing of hypothesis? What are their advantages and drawbacks over the parametric methods? 2+3=5

Contd.

8. *(a)* The following are the measurements of breaking strength of a certain kind of 2 inch cotton ribbon in pounds : 5

163, 165, 189, 161, 171, 158, 151, 169, 162, 163, 139, 172, 165, 148, 166, 172, 163, 187, 173.

Use sign test to test the null hypothesis $\mu = 160$ against the hypothesis $\mu > 160$ at the 0.05 level of significance.

Or

(b) A coaching institute claims that their coaching helps raising the SAT score. The SAT scores of students who took SAT test before and after coaching are as follows : 5

Before coaching :	980	1200	850	1100	150	175	145
After coaching :	1050	1120	950	1200	168	170	125

Use Wilcoxon signed rank test at 0.05 level of significance to test the coaching institute's claim.

---

1 SEM PG (CBCS) STS 200

2021

(held in February/March, 2022)

COMMERCE

Paper : 10200

(Statistics)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer ***any five*** of the following : 2×5=10

   (a) Define independent and mutually exclusive events. Can two events be mutually exclusive and independent simultaneously ?

   (b) What is the probability that a leap year selected at random will contain 53 Mondays ?

(c) Differentiate with example between simple linear regression model and multiple linear regression model.

(d) "Probability sampling has several advantages over judgement sampling." Discuss.

(e) Define the errors associated with testing of hypothesis.

(f) Distinguish between stratified random sampling and cluster sampling.

2. Answer ***any five*** of the following : 4×5=20

(a) A class consists of 30 students of which 20 are boys and 10 are girls. Half of the boys and half of the girls are blue-eyed. A person is selected at random. What is the probability that the person is a boy or blue eyed ?

(b) The chance that *A* tells the truth in narrating a particular incident is 70% and the chance that *B* tells the truth in narrating another incident is 80%. What is the probability that both of them contradict each other in describing a particular incident.

(c) Suppose it is desired to determine whether 10 students of class IX can perform equally in two different tests. The test scores are as follows :

Test 1 :	90	90	80	90	92	88	90	63	70	54
Test 2 :	84	84	82	94	90	85	89	62	65	52

[Tabulated value of t for 5% level of significance for 9 d.f. is 2·262]

(d) Two random samples drawn from two normal populations are as follows :

Sample I : 20 16 26 27 23 32 18 24 25 19

Sample II : 27 33 42 35 32 34 38 28 41 43 30 37

Test whether the population have same variances. [Tabulated value of F for (11, 9) d.f. at 5% level of significance is 3, 11].

(e) Find the line of regression of y on x from the following data :

x	5	10	15	25	30	35	40	45
y	25	32	44	32	39	49	55	60

What will be the value of y for $x = 48$?

(f) The following are the measurement of breaking strength of a certain kind of 2 inch cotton ribbon in pounds :

163, 165, 160, 189, 161, 171, 158, 151, 169, 162, 163, 139, 172, 165, 148, 166, 172, 163, 187, 173

Use sign test to test the null hypothesis $\mu = 160$ against the hypothesis $\mu > 160$ at the 0·05 level of significance.

(g) The table given below is the sexwise classification of a group of individuals based on their employment status :

	Male	**Female**	**Total**
Employed	50	20	70
Unemployment	15	15	30
Total	65	35	100

Test if the employment status have same relation with the sex of the individuals.

[Table value of χ^2 for 1 d.f. at 5% level of significance is 3·841].

M.Com 1st Semester Examination 2020 (held in June, 2021)

Course Title: Paper: Statistics

Course Number: 10200

Nature of the Course: Core

Full Marks: 30

Time: One and Half Hour

Mode of Examination: Online (Open Book Examination)

---

1. Answer the following questions **(Word Limit for 1(a), (b) and (c) is 100 each)**:

3X4=12

(a) In what respects sample survey is superior to a census?

(b) What is the difference between stratified random sampling and cluster sampling?

(c) Why we use non-parametric tests instead of using parametric tests?

(d) Let X be a random variable with the following probability distribution, find $E(3X+2)^2$.

X	-2	5	10
P(x)	1/3	1/6	1/2

2. Answer any two of the following questions:

2X4=8

(a) Dibrugarh University authority decide to form a committee of 4 people from 5 professors of Mathematics department, 4 professors of Economics department and 6 professors of Commerce department. Find the probability of forming the committee such that it should have at least one professor from the Commerce department.

(b) A bag contains 15 cards numbered with 1 to 15. A card was drawn at random and not replaced to the bag. After that similarly the second card was drawn. What is the probability that the first card was an even numbered card and the second was an odd numbered card?

(c) The first paper of UGC NET exam contains 60 questions. Each question has four possible answers, out of which only one is correct. Find the probability of getting 50 correct answers if students solve questions randomly.

3. Answer any two of the following questions: 5x2=10

(a) Obtain the line of regression for the following data.

Dependent Variable (Y): 67 68 65 68 72 72 69 71

Independent Variable(X): 65 66 67 67 68 69 70 72

(b) The results of a survey regarding radio listeners' preference for different type of music are given below,

Type of Music Preferred	**Age Group**		
	19-30	**Above 30**	**Total**
National Music	20	80	100
Foreign Music	150	50	200
Total	170	130	300

Is preference of type of music influenced by age? ($\chi^2_{0.05,\,1} = 3.841$)

(c) According to the company the median life (in hours) of a 25 watt bulb is 200 hours. But a shopkeeper claims that it is more than 200 hours. A random sample of 12 bulbs yielded the following results: 220, 222, 240, 230, 212, 192, 188, 220, 190, 210, 208, and 194. Use the Wilcoxon signed rank test with 5% level of significance to test the shopkeepers claim. (Critical value for n=12 and α=0.05 for a single tailed test is = 17)

****

1 SEM MCOM (CBCS) STS 200

2022

(December)

COMMERCE

Paper : 10200

(Statistics)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Select the correct answer : 1×10=10

(a) Mutually exclusive events are always—

(i) Independent

(ii) Dependent

(iii) Equally likely

(iv) None of the above

(b) If S is a certain event and ϕ is an impossible event then

(i) $P(S \cup \phi) = P(S)$

(ii) $P(S \cup \phi) = P(\phi)$

(iii) $P(S) = P(\phi)$

(iv) All of the above

(c) The conditional probability $P(A/B)$ will be equal to $P(A)$ when A and B are—

(i) Mutually exclusive events

(ii) Exhaustive events

(iii) $P(A \cup B) = \phi$

(iv) Independent events

(d) In case of Binomial Distribution —

(i) Mean > Variance

(ii) Mean < Variance

(iii) Mean = Variance

(iv) None of the above

(e) If A and B are two independent events, then

(i) $\overline{A}$ and $\overline{B}$ are also independent

(ii) A and $\overline{B}$ are also independent

(iii) $\overline{A}$ and B are also independent

(iv) All of the above

(f) Which of the following is a non-probability sampling technique?

(i) Simple Random Sampling

(ii) Cluster Sampling

(iii) Stratified Random Sampling

(iv) Systematic Sampling

(g) Non-parametric tests are

(i) distribution free

(ii) distribution free but non-random

(iii) population followers normal distribution

(iv) same as parametric tests

(h) If R is coefficient of determination, then

(i) $R \geq 0$

(ii) $R \leq -1$

(iii) $R \geq 0,\ R \leq -1$

(iv) $0 \leq R \leq 1$

Contd.

(d) What is conditional probability? Explain with example.

(e) Differentiate between simple linear regression model and multiple linear regression model with example.

(f) Distinguish between stratified random sampling and cluster sampling.

(g) Find the probability that a leap year selected at random will contain 53 Sundays.

(h) If X is a random variable and C is any constant, then prove that

$$V(aX) = a^2V(X)$$

3. (a) What is random variable? Write down the axioms of the probability. 1+3=4

Or

(b) Write briefly on SPSS package. 4

4. Write short notes on : ***(any four)*** 4×4=16

(a) Probability and Non-probability sampling

(b) Complete enumeration and sample survey

(c) An Ice-cream company wants to know if there is any association between the gender and choice of ice-cream flavours. The company collected the data from 200 samples which are summarized in the following table. Test if the choice of ice-cream flavour is related to the gender.

$\left(\chi^2_{1,\,0.05} = 3.841\right)$

Flavour→ / Gender ↓	Chocolate	Strawberry
M	75	30
F	45	50

400

1 SEM MCOM (CBCS) STS 200

2023

(December)

COMMERCE

Paper : 10200

(Statistics)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Select the correct answer : 1×10=10

(a) Probability can take values

(i) $-\infty$ to ∞

(ii) $-\infty$ to 1

(iii) –1 to 1

(iv) 0 to 1

(b) Two events are said to be independent if :

(i) each outcome has an equal chance of occurrence

(ii) there is no common point in between them

(iii) one does not affect the occurrence of the other

(iv) both the events have only one point

(c) Two random variables X and Y is said to be independent if :

(i) $E(XY) = 1$

(ii) $E(XY) = 0$

(iii) $E(XY) = E(X)\,E(Y)$

(iv) $E(XY)$ = any constant value

(d) The height of persons in a country is a random variable of the type :

(i) continuous random variable

(ii) discrete random variable

(iii) neither discrete nor continuous random variable

(iv) continuous as well as discrete random variable

(e) A family of parametric distribution in which mean is equal to variance is :

(i) binomial distribution

(ii) uniform distribution

(iii) normal distribution

(iv) Poisson distribution

(f) A family of parametric distribution in which mean is always greater than its variance is :

(i) binomial distribution

(ii) geometric distribution

(iii) Both (i) and (ii)

(iv) Neither (i) and (ii)

(g) Which of the following is a non-probability sampling technique ?

(i) Simple random sampling

(ii) Stratified random sampling

(iii) Purposive sampling

(iv) Cluster sampling

(h) Probability of drawing a unit at each selection remains same.

(i) Simple random sampling without replacement

(ii) Simple random sampling with replacement

(iii) Both (i) and (ii)

(iv) None of (i) and (ii)

(i) The hypothesis under the test is known as :

(i) simple hypothesis

(ii) alternative hypothesis

(iii) null hypothesis

(iv) None of the above

(j) A wrong decision about H_0 leads to :

(i) one kind of error

(ii) two kinds of error

(iii) three kinds of error

(iv) four kinds of error

2. Answer **any five** of the following : 2×5=10

(a) Give the classical definition of probability.

(b) Narrate addition theorem of probability.

(c) What are probability mass function and probability density function ?

(d) Differentiate between simple linear regression model and multiple linear regression model with examples.

(e) Distinguish between complete enumeration and sampling survey.

(f) Explain probability and non-probability sampling.

(g) Define null and alternative hypotheses.

3. Answer **any five** of the following : 4×5=20

(a) What are the principal steps involved in hypothesis testing of a statistical test ?

(b) Explain stratified random sampling and cluster sampling.

(c) Delineate the main steps of sampling design.

(d) Delineate binomial distribution and its important properties.

(e) Discuss elaborately normal distribution and its characteristics.

(f) State the conditional probability and Bayes' theorem.

(g) Describe briefly ordinary sign test.

4. Solve ***any two*** of the following : 5×2=10

(a) Two dice are thrown simultaneously. Find the probability of getting even number on both the dice.

(b) The probability that a contractor will get a plumbing contract is $\frac{2}{3}$, and the probability that he will get an electric contract is $\frac{5}{9}$. If the probability of getting *at least one* contract is $\frac{4}{5}$, what is the probability that he will get both the contracts ?

(c) A coin is tossed six times. What is the probability of obtaining four or more heads ?

(d) In a bolt factory machines *A*, *B*, *C* manufacture respectively, 25%, 35% and 40% of the total production. Of there output 5%, 4% and 2% respectively are defective bolts. A bolt is drawn at random and is found to be defective. What is the probability that it is manufactured by machine *A* ?

Write short notes on : ***(any two)*** 5×2=10

(i) Heteroscedasticity

(ii) Multicollinearity

(iii) Type I and Type II errors

(iv) Non-parametric statistical methods

---

1 SEM MCOM (CBCS) STS 200

2024

(December)

COMMERCE

Paper : 10200

(Statistics)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Select the correct answer : 1×10=10

(a) Classical probability is possible in case of :

(i) Unequilikely outcomes

(ii) Equilikely outcomes

(iii) Either Unequilikely or equilikely outcomes

(iv) All of the above

(b) If $P(A|B)=\frac{1}{4}$ and $P(B|A)=\frac{1}{3}$ then $\frac{P(A)}{P(B)}$ is equal to

(i) $\frac{3}{4}$

(ii) $\frac{7}{12}$

(iii) $\frac{4}{3}$

(iv) $\frac{1}{12}$

(c) A family of parametric distribution in which the mean is always greater than its variance is :

(i) Binomial distribution

(ii) Geometric distribution

(iii) Both (i) and (ii)

(iv) Neither (i) nor (ii)

(d) What is the significance of randomization in sampling methods ?

(i) It ensures fairness

(ii) It eliminates all biases

(iii) It helps to generalize findings to the population

(iv) All of the above

(e) What type of sampling involves selecting every n^{th} individual ?

(i) Stratified sampling

(ii) Simple random sampling

(iii) Systematic sampling

(iv) Cluster sampling

(f) Which of the following assumptions is not required for multiple regression analysis ?

(i) Linearity

(ii) Homoscedasticity

(iii) Multicollinearity

(iv) Independence of observations

(g) What does the R-squared value indicate in multiple regression analysis ?

(i) Strength of the relationship

(ii) Direction of the relationship

(iii) Number of predictors

(iv) Significance of coefficients

(h) Which of the following represents a Type I error ?

(i) Rejecting H_0, when it is true

(ii) Failing to reject H_0, when it is false

(iii) Accepting the H_1, when it is false

(iv) None of the above

Contd.

(i) What is the main assumption of parametric tests ?

(i) Data is normally distributed

(ii) Data is skewed

(iii) Sample size is small

(iv) Data is nominal

(j) What is the primary purpose of using SPSS in research ?

(i) Data entry

(ii) Data analysis

(iii) Data collection

(iv) Data visualization

2. Answer **any five** of the following : 2×5=10

(a) Explain the concept of sampling.

(b) The probability that a contractor will get a plumbing contract is $\frac{2}{3}$, and the probability that he will get an electric contract is $\frac{5}{9}$. If the probability of getting at least one contract is $\frac{4}{5}$, what is the probability that he will get both the contracts ?

(c) Differentiate between the terms 'statistic' and 'parameter'.

(d) What are the assumptions of multiple regression ?

(e) What is level of significance ?

(f) State the meaning of the null and alternate hypothesis.

(g) What is the difference between MS Excel and MS Access ?

3. Answer **any five** of the following : 4×5=20

(a) Write a short note on : Bayes' theorem.

(b) Jio AirFiber is planning to add one new channel to its basic service. There are about 20,000 subscribers, and the company knows that 35 per cent of these are college students, 45 per cent are white-collar workers, 15 per cent are blue-collar workers, and 5 per cent are other. However, the company believes there is much variation within these groups. Briefly explain, which of the following sampling methods is more appropriate for this case: random, systematic, stratified, or cluster sampling.

(c) Two bags, one containing 5 white and 3 black balls and the other contains 4 white and 5 black balls. Two balls are drawn at random from any one of them. What is the probability that two selected balls are white ?

(d) What do you understand by the coefficient of determination?

(e) Discuss the steps of hypothesis testing.

(f) Explain the different types of errors in hypothesis testing.

(g) Explain the procedure for conducting an F-Test.

(h) Explain the significance of data processing in statistical analysis.

4. Answer ***any two*** of the following: 10×2=20

(a) Indian Cinema Ltd. knows that a popular movie ran an average of 90 days in each city across its theaters nationwide, with a standard deviation of 12 days. The manager of the South Indian region is interested in comparing the movie's popularity in his region with that in all other regions. He randomly selected 50 theaters in his region and found that they ran the movie for an average of 85 days.

(i) State appropriate hypotheses for testing whether there is a significant difference in the length of the movie's run between theaters in the South Indian region and all of Indian Cinema Ltd.'s other theaters.

(ii) At a 5% significance level, test the framed hypotheses. 2+8=10

(b) Bulbuli Restaurant, is concerned about the sales behaviour of roti-sabji sold at the store. The owner realizes that there are many factors that might help explain sales, but believes that number of sabji items and price are major determinants. The owner collected the following data :

Number of Plates Sold	Number of Sabji Items	Price
30	3	30
60	6	20
70	10	50
80	12	40
22	2	55
25	6	50

(i) Calculate the least-squares equation to predict the number of plates sold based on the number of sabji items and their price.

(ii) If sabji items are 15 and price is 42, what sales (number of plates) would you predict? 8+2=10

Contd.

least one pr... ...nds numbered with 1 to 15. ... ...nd ca...

(c) Solve the following : 5+5=10

(i) Two bags, one containing 5 white and 3 black balls and the other contains 4 white and 5 black balls. Two balls are drawn at random from any one of them. What is the probability that two selected balls are white ?

(ii) A machine produces 5% defective screws. What is the probability that a batch of 15 screws contains not more than 2 defective screws ?

(d) Write short notes on : 5+5=10

(i) SPSS package

(ii) MS Excel