B.Sc. (Honours) Semester II Examination 2022
Subject: Statistics
Paper: CC 3
Probability and Probability Distribution (Theory)

Full Marks: 40 Time: 3 hours.

Answer any four of the following six questions of equal marks.
(Notations carry usual meanings)

1. (a) r balls are distributed into n cells following the exclusion principle. Find the probability that exactly one cell will remain empty.
   (b) Is it true that $Cor(x, y) = 0$ implies X and Y are independent? Answer with relevant example. What do you think about the reverse?

2. (a) Obtain moment generating function of $Bin(n, p)$ distribution and hence find its mean and variance.
   (b) A random variable X has the density function
   $$f(x) = \frac{1}{2\sqrt{x}}; 0 < x < 1$$
   Obtain the moment generating function of $Y = 7X - 2$. Hence find first cumulant of Y.

3. (a) Given A and B are two independent events defined on a sample space. Prove that A^C and B^C are also independent
   (b) Prove that r^{th} order central moment can be deduced from r^{th} and lower order raw moments.

4. (a) Let X be a discrete random variable with probability mass function
   $$f(x) = \frac{1}{n+1}; x = 0, 1, 2, \ldots, n$$
   Find $E(X)$ and $Var(X)$.
   (b) Prove that for k random variables $X_1, X_2, \ldots, X_k$ and arbitrary constants $a_1, a_2, \ldots, a_k$
   $$Var(\sum_{i=1}^{k} a_i X_i) = \sum_{i=1}^{k} a_i^2 Var(X_i) + 2 \sum \sum_{i<j} a_i a_j Cov(X_i, X_j)$$

5. (a) State and prove Baye's theorem.
   (b) Let the joint probability density function of two dimensional random variable (X, Y) be
   $$f(x, y) = 2(x + y - 3xy^2); 0 < x < 1; 0 < y < 1$$
   In this case is $E(XY) = E(X)E(Y)$? Also find $E(X/Y = y)$.

6. (a) Show that Poisson distribution can be looked upon as a limiting form of the binomial distribution. (State the assumptions you need).
   (b) Suppose that $P(\geq x)$ is given below for a random variable X
   $$P[X \geq x] = \frac{3}{(1+x)^2} - \frac{2}{(1+x)^3}; x > 0$$
   Find the corresponding density function and the value of the constant c such that $P(X > c) = \frac{1}{2}$.

B.Sc. (Honours) Examination 2022
Semester II
Statistics
Paper: CC 3B
Probability and Probability Distribution (Practical)

Full Marks: 20 Time: 2 hours.

Answer all the questions
(Notations carry usual meanings)

1. An insurance company insures 4000 people against loss of both eyes in a car accident. Based on previous data, the rates were computed on the assumption that on the average 10 persons in 10000 will have car accident each year that result in this type of injury. What is the probability that more than 3 of the insured will collect on their policy in a given year? 4

2. In an intelligent test administered to 1000 children the average score is 42 and standard deviation 24. The score is normally distributed.

   (a) Find the number of children exceeding the score 60.

   (b) Find the number of children with score lying between 20 and 40.

   4

3. Seven coins are tossed and number of heads noted. The experiment is repeated 128 times. The distribution is obtained.

Number of heads	0	1	2	3	4	5	6	7	**Total**
Frequency	7	6	19	35	30	23	7	1	128

   Fit a binomial distribution assuming the coin is unbiased. 6

4. A random variable X has the following probability distribution.

X=x	0	1	2	3	4	5	6	7	8
P[X=x]	k	3k	5k	7k	9k	11k	13k	15k	17k

   (a) Determine the value of k

   (b) Find $P(X < 3)$ and $P[X = 2/X < 5]$

   3

5. Viva-voce. 3

B.Sc. Examination, 2022
Semester-II
Statistics
Course: CC-4
(Algebra)

Time: 3 Hours **Full Marks: 60**

Questions are of value as indicated in the margin.
Notations have their usual meanings

Group-A (Answer any five questions)

1. Answer any five questions. [5 × 2]

(a) Illustrate whether matrix product is always commutative or not.

(b) Define Subspace and span. When do we say a set of vectors are linearly independent?

(c) Find reduced row-echelon form of $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$.

(d) Let $A_{m\times n}$ be a real matrix. Show that rank of A is same as rank of $A^T A$.

(e) Illustrate what will be the degree of the polynomial obtained by sum of two different polynomials of degree m and n? What will be the degree of the polynomial obtained by product of two different polynomials of degree m and n?

(f) Define determinant of a matrix.

(g) Define symmetric matrix and skew-symmetric matrix.

(h) Show that, "similarity" of matrices is an equivalence relation. In other words, show that, "similarity" is reflexive, symmetric and transitive relation.

Group-B (Answer any five questions) [5 × 10]

2. (a) A is a 3×3 non-null real matrix and $A^2 - A - I_3$ is a null matrix. Show that A^{-1} exists and $A^{-1} = A - I_3$.

(b) If A and B are two $n \times n$ matrices and A has an inverse, then show that

$$(A+B)A^{-1}(A-B) = (A-B)A^{-1}(A+B)$$

(c) Investigate for what values of a, b the following system of equations

$$\begin{aligned} x + y + z &= 1 \\ x + 2y - z &= b \\ 5x + 7y + az &= b^2 \end{aligned}$$

has only one solution, no solution or infinitely many solutions.

[2 + 4 + 4]

3. (a) Show that, rank of an orthogonal matrix of order n is n.

(b) Show that, the number of orthogonal diagonal matrices of size $n \times n$ is 2^n.

(c) If $(I + A)^{-1}(I - A)$ is skew-symmetric matrix, then show that A is a real orthogonal matrix.

[2 + 4 + 4]

CONTINUED

4. (a) When do we say a symmetric matrix is in "normal form"?

(b) Define index and signature of a matrix in normal form.

(c) Reduce the real quadratic form $2x^2 + 2y^2 + 5z^2 - 4xy - 2xz + 2yz$ to its normal form and find its rank and signature.

.

[2 + 4 + 4]

5. (a) Define idempotent matrix and orthogonal matrix.

(b) Prove that, if A and B are two matrices such that $AB = A$ and $BA = B$ then A, B are idempotent.

(c) Let A is an idempotent matrix. Show that, all eigen values of A are either 0 or 1.

[2 + 4 + 4]

6. (a) Define Similar matrix with an example.

(b) If A, B are two square matrices of same order and B^{-1} exists, then show that A and $B^{-1}AB$ have have same eigenvalues.

(c) If $A = \begin{bmatrix} a^2 & 2 & 2 & 2 \\ 4 & b^2 & 3 & 3 \\ 6 & 6 & c^2 & 5 \\ 7 & 7 & 7 & 7 \end{bmatrix}$, show that A is nonsingular matrix, for $a, b, c \in Q$, where Q is the set of rational numbers. .

[2 + 4 + 4]

7. (a) Define eigenvalue and eigenvector of a square matrix.

(b) Find eigenvalues of the following matrix. Also find the algebraic multiplicity of each of the eigenvalues.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 3 & -1 \end{bmatrix}$$

(c) Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Find the characteristic polynomial. Hence, or otherwise, Show that $A^2 - 5A - 2I = 0$.

.

[2 + 4 + 4]

B.Sc. (Honours) Examination, 2022
Semester-IV
Statistics
Course: CC-8A (Survey Sampling and Indian Official Statistics)
Full Marks: 40 Time: 3 Hours

1. Answer the following questions. $1 \times 10 = 10$

   (a) Write down the full form of DGCIS.
   (b) Fill in the blank: NSSO was started in the year ____.
   (c) How many divisions does the NSSO have?
   (d) Fill in the blank: NSSO was headed by ____.
   (e) What are the two wings of MoSPI?
   (f) Fill in the blank: 'The Statistical Abstract— India' is published by ____.
   (g) Mention one major responsibility of NAD.
   (h) In which year the first census was taken?
   (i) Mention one important publication from the Social Statistics Division.
   (j) Define index of industrial production.

2. (a) What do you mean by linear systematic sampling? Why is it called a mixed sampling scheme?
   (b) Show that in the presence of a linear trend in the population, the linear systematic sampling provides a more efficient estimator of the population total than the simple random sampling.

   (2+2)+6=10

3. (a) In the context of Stratified random sampling, explain how can you determine the optimal sizes of the samples drawn from each startum.
   (b) In this context, prove that $V_{ran} \geq V_{prop} \geq V_{opt}$, where the symbols have their usual meanings.

   5+5=10

4. (a) A simple random sample of size $n = n_1 + n_2$ with mean $\overline{y}$ is drawn from a finite population and a simple random subsample of size n_1 is drawn from it with mean $\overline{y_1}$. Show that
      i. $V\left(\overline{y_1} - \overline{y_2}\right) = {S_y}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$, where $\overline{y_2}$ is the mean of the remaining n_2 units in the sample.
      ii. $V\left(\overline{y_1} - \overline{y}\right) = {S_y}^2 \left(\frac{1}{n_1} - \frac{1}{n}\right)$
      iii. $Cov\left(\overline{y}, \overline{y_1} - \overline{y}\right) = 0$
   (b) Prove that under simple random sampling without replacement, the sample variance with divisor $n-1$, i.e. $\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \overline{y})^2$ is an unbiased estimator of the population variance $\frac{1}{(N-1)}\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2$

   6+4=10

5. (a) Define the ratio estimator of the population mean. Mention the situation when it is appropriate to use.
   (b) Find the expressions of the approximate bias and MSE of the estimator.
   (c) Find the condition under which the ratio estimator performs better than the usual SRS estimator of the population mean.

2+6+2=10

6. (a) Compare the relative merits and demerits of cluster sampling and two-stage sampling.
   (b) Suggest an unbiased estimator of population total under two-stage sampling. Find its variance and an unbiased estimator of the variance,

2+(1+7)=10

B.Sc. (Honours) Examination, 2022
Semester-IV
Statistics (Practical)
Course: CC-8B (Survey Sampling and Indian Official Statistics (Practical))
Full Marks: 20 Time: 2 Hours

(1) Select 10 random points within the ellipse $9x^2 + 16y^2 = 144$. 5

(2) The following is the population of B.Sc. (Statistics) students of a certain university, together with the data on two variables, viz. Marks in Sample survey and Marks in Estimation (both out of 50).

Serial Number	Name of the student	Marks in Sample Survey	Marks in Estimation
1	Aneek Sarkar	32	34
2	Rajdeep Nandi	45	38
3	Snehangshu Bhuin	37	42
4	Papu Mahanta	27	35
5	Krishnendu Dutta	48	42
6	Debapriyo Sengupta	40	40
7	Md. Inamul Haque	33	39
8	Sarnalina Datta	41	46
9	Titir Mandal	28	34
10	Anindita Kundu	37	34
11	Yash Prasad	39	36
12	Ananya Bakshi	29	28
13	Sandip Mishra	34	31
14	Sujash Krishna Basak	40	45
15	Rohit Jana	38	47
16	Pabitra Mondal	43	35
17	Aritraa Karmakar	32	41
18	Aritra Ghosh	41	37
19	Bidisha Bhandari	40	42
20	Sumit Kumar	34	38

(a) Draw 5 distinct students at random from the population.
(b) Estimate the average marks in Estimation. Also find the standard error of your estimate, based on your sample.
(c) Obtain the ratio and regression estimate of the average marks in Estimation, taking the marks in Sample Survey as the auxiliary information.
(d) Find the estimated mean square error of the ratio and regression estimators.
(e) Comment on the performances of the estimators. 2+2+3+3+2=12

(3) Practical Note Book and Viva-voce. 3

B.Sc. (Honours) Examination, 2022
Semester- IV
Statistics
Course: CC-9
(Linear Models [Theory])
Time: 3 Hours   Full Marks: 40

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer any four questions

1. a) State and prove Gauss-Markov theorem.
b) It is given that y_1, y_2, y_3 are independent with same variance and $E(y_1) = E(y_3) = \beta_1 + \beta_2$, $E(y_2) = \beta_1 + \beta_3$. Show that $\sum_{i=1}^{3} c_i\beta_i$ is estimable if $c_1 = c_2 + c_3$. 6+4

2. a) What do you mean by linear zero function? Show that a linear function in $y = [y_1, \cdots, y_n]'$ is the BLUE of its expectation if and only if it is uncorrelated with every linear zero function.
b) In usual notations, show that $R(X) = R(X'X) = R(XX')$. (1+6)+3

3. Write down detailed analysis, assumption and hypothesis testing of a two way ANOVA with m observations per cell. 10

4. a) Briefly write down the steps followed in regression analysis.
b) What do you mean by centered model? What are the OLS estimators of its two parameters? Are they unbiased? What is the covariance between them? Find variances of these two estimators. 3+7

5. a) Find an unbiased estimator of σ^2 for a simple linear regression model: $y = \beta_0 + \beta_1 x + \varepsilon$.
b) In usual notations for simple linear regression, show that: i) $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i$, ii) $cov(\hat{y}, e) = 0$, iii) $[cor(y, \hat{y})]^2 = [cor(x, y)]^2$. 4+(2+2+2)

6. a) Find the least square estimator of $\beta = [\beta_0, \beta_1, \beta_2, \cdots, \beta_k]'$ for the multiple linear regression model: $y = X\beta + \varepsilon$, where $y = [y_1, \cdots, y_n]'$, $\varepsilon = [\varepsilon_1, \cdots, \varepsilon_n]'$ and X is a matrix of order $n \times p$.
b) In usual notations, show that $V(\hat{y}) = \sigma^2 H$. 6+4

B.Sc. (Honours) Examination, 2022
Semester-IV
Statistics
Course: CC-9B
(Practical on Linear Models)
Time: 2 Hours **Full Marks: 20**

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **all** questions

1. The purity of oxygen produced by a fraction process is thought to be related to the percentage of hydrocarbons the main condenser of the processing unit. Fourteen samples are given.

PURITY (%)	HYDROCARBON (%)	PURITY (%)	HYDROCARBON (%)
86.91	1.02	96.73	1.46
89.85	1.11	99.42	1.55
90.28	1.43	98.66	1.55
86.34	1.11	96.07	1.55
92.58	1.01	93.65	1.40
87.33	0.95	87.31	1.15
86.29	1.11	95.00	1.01

a) Fit a simple linear regression model to the data
b) Test the hypothesis $H_0 : \beta_1 = 0$.
c) Calculate R^2.
d) Find a 95% confidence interval on the slope.
e) Write down the ANOVA table and discuss. 2+2+2+2+2

2. An experiment was conducted to determine the effects of 4 different varieties of cowpeas(V1,V2,V3,V4) & 3 different spacing's (S1,S2,S3) and also see if the varieties behave differently at different spacing's. The data given below give the yield of each of the 3 plots taken for each variety. Spacing combination

Spacing									
Variety	S1			S2			S3		
V1	45	43	56	60	50	45	66	57	50
V2	58	55	61	60	59	54	59	55	51
V3	53	49	63	65	56	50	66	58	52
V4	61	60	65	60	58	56	53	53	48

Carry out an ANOVA for the above data. 7

3. Practical note book and viva-voce. 3

B.Sc (Honours) Examination, 2022

Semester- IV

Statistics

Course: CC-10

(Statistical Quality Control)

Time: 3 hours **Full Marks: 40**

Questions are of values as indicated in the margin

Answer **any four** questions

1. A. Distinguish between:
    i) Process control and Product Control
    ii) Control limits and Specification limits

   B. Describe, in detail, the construction of control chart for number of defects. Mention a few application area where this chart can be used. 4+6

2. A. Describe the construction of $(\bar{X}, R)$ chart.

   B. Describe the technique of sampling inspection by variables in the normal distribution case. 5+5

3. A. When a system is said to be out of control?

   B. How would you modify a control chart if sample sizes are varying? Give an example. 3+7

4. A. Define the following

   Rational subgroup, process capability

   B. Describe the construction of control chart for number of defectives. 4+6

5. A. Define the following

   LTPD, AOQ, AOQ, ASN and ATI

   B. For a double sampling inspection plan, find expressions for all the terms mentioned in (A). 3+7

6. A. Define Consumer’s risk and producer’s risk.

   B. Assuming a binomial distribution find the expression for the producer’s risk, consumer’s risk, AOQ and ASN functions for a single sampling inspection plan. 3+7

BSC Semester IV Examination, 2022
Statistics
Paper: [CC-10 B] Statistical Quality Control (Practical)

Full Marks: 20 Time: 2 hours

1. Assume that in the manufacture of 1 kg Mischmetal ingots, the product weight varies with the batch. Below are a number of subsets taken at normal operating conditions, with the weight values given in kg. Construct the (X-bar, R) chart on the basis of these 11 subsets. Measurements are taken sequentially in increasing subset number. Also comment on your findings.

Subset #	Values (kg)
1 (control)	1.02, 1.03, 0.98, 0.99
2 (control)	0.96, 1.01, 1.02, 1.01
3 (control)	0.99, 1.02, 1.03, 0.98
4 (control)	0.96, 0.97, 1.02, 0.98
5 (control)	1.03, 1.04, 0.95, 1.00
6 (control)	0.99, 0.99, 1.00, 0.97
7 (control)	1.02, 0.98, 1.01, 1.02
8 (experimental)	1.02, 0.99, 1.01, 0.99
9 (experimental)	1.01, 0.99, 0.97, 1.03
10 (experimental)	1.02, 0.98, 0.99, 1.00
11 (experimental)	0.98, 0.97, 1.02, 1.03

5

2. Following are the figures for the number of defectives in 16 lots, each containing 2,000 rubber belts:
(341-your roll number), 225, 322, 280, 306, 337, 305, 356, 402, 216, 264, 126, 409, 193, 326, (280+your roll number)
Drawing the control chart for fraction defective, plot the points on it. Comment on the state of control of the process. 5

3. A. Suppose a tyre supplier ships tyres in lots of size 400 to the buyer. A single sampling plan with n = 15 and c = 0 is being used for the lot inspection. The supplier and the buyer's quality control inspector decide that AQL = 0.01 and LTPD = 0.10. Compute the producer's risk and consumer's risk for this single sampling plan.
   B. Suppose the rejected lots are screened and all defective tyres are replaced by non-defective tyres. Construct the AQO curve for this plan. 4+3

4. Practical note book and Viva-voce 3

B.Sc. (Honours) Examination, 2022
Semester-IV
Statistics
Course: SECC-2
(Statistical Data Analysis Using Software Packages)

Time: 2 Hours **Full Marks: 25**

Questions are of value as indicated in the margin
Notations have their usual meanings

Group – A (Answer any ten questions) $10 \times 1 = 10$

1. Answer the following questions with proper justification.

   (a) In python, how do you find $\binom{10}{4} + \binom{10}{8}$?

   (b) How to use stem-leaf plot in python?

   (c) Round off $\sqrt{7}$ upto 4 decimal point?

   (d) Which python library enable us to use dataframe?

   (e) In which situation we use boxplot? Create a data of length 10 and represent it graphically by boxplot in python.

   (f) Using python, find the median for the data (x): 45 26 19 49 43 with respective frequencies (f): 2 4 3 5 4.

   (g) Find the raw moments upto order 4 for the data (x): 60 39 75 33 69 47 69 32 59 52 49 63 using python.

   (h) In python, generate a sqeuence from 1 to 50 with increment 0.2.

   (i) In usual notations, generate 50 random samples from negative binomial distribution with parameters $r = 2,\ p = 0.3$.

   (j) Write down five different colour code in python.

   (k) Write down a measure of kurtosis. Create a hypothetical data of length 8 and compute your mentioned measure using python.

   (l) Using python draw the graph of $y = e^x$ for $x \in [-3, 3]$.

   (m) In python, how you can increase the size of $x - label$ of a graph?

Group – B (Answer any three questions) $3 \times 5 = 15$

2. Using python, find the range, standard deviation, mean deviation about mean, inter-quartile range and coefficient of variation for the data (x): 16 40 12 46 47 64 50 64 56 37 82 78 27 13 36 50 51 75 70 87. 5

3. Suppose we have a bivariate data of 20 students, Marks in School (x): 148 134 131 146 135 159 136 161 166 165 160 136 176 145 144 153 158 145 126 170 and Marks in College (y): 423 405 362 369 333 417 301 425 372 438 415 393 349 306 380 338 450 326 381 359. Using python, plot a scatter diagram, correlation coefficient between x, y and correlation matrix. 5

4. Using python, plot the multiple graphs of normal distribution from $N(0.7, 1.2)$, $N(0.7, 0.6)$, $N(0.7, 1.8)$ and $N(0.7, 2.4)$. 5

5. Suppose there are 90 people numbered from 1 to 90. Use circular systematic sampling to select 17 lucky people who will get movie ticket. Explain the process briefly together with the python code. 5

6. For a Poisson distribution i.e. $Poi(\lambda)$, it is given that $\overline{x}$ is an estimator of λ. Now with the help of Python, generate a random sample of size 100 from $Poi(6)$ and using that sample find $\hat{\lambda}$. Repeat the process for sample size 500 and 5000. Write down your observation. 5

B.Sc. (Honours) Examination, 2022
Semester-VI
Statistics
Course: CC-13A
(Design of Experiments)
Time: Three Hours Full Marks: 40

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **any five** questions

1. What is design of experiments? Discuss its principles. Discuss the impact of size and shape of plots and blocks on the result of an experiment? 4+4

2. What is a Latin Square Design (LSD)? How is it improvement over Randomized Block Design (RBD)? Discuss the layout and analysis of LSD. 1+2+2+3

3. How will you estimate the yield of a missing plot in an RBD when one observation is missing? Discuss in detail the analysis of an RBD after estimating the yield of missing plot. 3+5

4. Distinguish between the ANOVA model and the ANOCOVA model for RBD with one concomitant variable. How do you judge whether the inclusion of the concomitant variable is worthwhile or not? If worthwhile, give the detailed analysis. 2+2+4

5. Define the terms main effects and interaction effects in relation to a 2^3experiment and show that they are mutually orthogonal. How do you obtain the SS due to main effects or interaction effects in a 2^3 experiment? Give only the AVOVA table of a 2^3 experiment conducted in randomized blocks. 4+2+2

6. Consider (2^4, 2^2) design. Treatment combinations belonging to a block is a, b, cd, abcd. (i) Construct the other three blocks. (ii) Find the confounded effects. (iii) Give the analysis of this experiment. 2+3+3

7. Differentiate the strip-plot experiment from the split-plot experiment. Discuss the layout and analysis of a strip-plot experiment in RBD. 3+5

8. Write short note on any two of the following: 4+4

(i) Confounding and local control principle in design of experiments.
(ii) Design of experiments over time and space.
(iii) Comparison of efficiency of LSD with respect to RBD.

B.Sc. (Honours) Examination, 2022
Semester-VI
Statistics
Course: CC-13B
(Practical on Design of Experiments)
Time: Two Hours Full Marks: 20

Questions are of value as indicated in the margin

1. Given below is the layout plan and wheat yield in kg per plot for a randomized block design with 5 treatments A, B, C, D, E.

Block:					
1	E2.50	D2.27	C1.62	B1.82	A0.91
2	C1.41	A0.95	E2.27	B1.95	D MISSING
3	A0.77	B2.04	E2.40	C1.82	D2.50

Obtain an estimate of the missing value and analyze the data. 7

2. The following table gives the plan and yields of an experiment involving three fertilizers N, P, K, each at two levels in eight blocks of four plots each.

Plan and yields of a 2^3-factorial experiment in blocks of four plots.

Block		Block	
Treatment	Yield	Treatment	Yield
(1)	145	k	189
pk	191	p	272
nk	300	n	160
np	240	npk	305

Block		Block	
Treatment	Yield	Treatment	Yield
(1)	226	p	226
k	159	nk	300
npk	240	pk	233
np	182	n	278

Block		Block	
Treatment	Yield	Treatment	Yield
p	186	n	209
npk	173	k	93
(1)	173	pk	224
nk	213	np	201

Block		Block	
Treatment	Yield	Treatment	Yield
pk	182	k	293
(1)	175	nk	226
npk	156	np	248
n	183	p	269

Analyze the data and write a report. 10

3. Practical Note Book and Viva-voce. 3

B.Sc. (Honours) Examination, 2022
Semester-VI
Statistics
Course: CC-14A
Multivariate Analysis & Nonparametric Methods
Time: 3 hrs Full Marks:40

Answer any **four** questions of the following.

1. (a) Show that multiple correlation is the highest possible correlation between X_1 and any linear combination on the set of independent variables $(X_1, X_2 \cdots, X_p)$.

   (b) Suppose you have a data set of size 8. You want to test if the central value of the population, the data coming from, is 18. Under this null hypothesis, deduce the expectation and variance of Wilcoxon signed rank test statistic.

   6+4

2. (a) Suppose you have a data set of size 9. Discuss a test procedure briefly to check if the data comes from standard normal variable.

   (b) Is the following model a linear regression model?X_1 and X_2 are independent variables while y being the dependent one.

   $$y = e^{x_1}\beta_1 + x_2\beta_2$$

   (c) When do you use Run test? Briefly state the rejection rules for run test under different alternatives.

   4+2+4

3. (a) Establish the relationship $\rho_{12.34\cdots p} = \frac{-\sigma^{12}}{\sqrt{\sigma^{11}\sigma^{22}}}$ where $\rho_{12.34\cdots p}$ being the population partial correlation coefficient and σ^{ij} being the (i,j) element of Σ^{-1}.

   (b) For a linear regression model coefficient of determination is 35%. Explain it.

   (c) Let $r_{12.3} = 0$, $r_{13.2} = 0$ and $r_{23.1} = 0$. Does it imply $r_{1.23} = 0$? Explain.

   6+2+2

4. (a) Let $X_p \sim N_p(\mu, \Sigma)$. Let us partition $\mathbf{X}_{p\times 1} = \begin{pmatrix} \mathbf{X}_{(1)_{2\times 1}} \\ \mathbf{X}_{(2)_{p-2\times 1}} \end{pmatrix}$ and $\mu = \begin{pmatrix} \mu_{(1)_{2\times 1}} \\ \mu_{(1)_{p-2\times 1}} \end{pmatrix}$ and $\Sigma_{p\times p} = \begin{pmatrix} \Sigma_{11 2\times 2} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Show that the conditional distribution of $X_{(2)}$ given $X_{(1)}$ follow a multivariate normal of order $p-2$.

   (b) Let $X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3 \left[(1,-1,3)', \Sigma = \begin{pmatrix} 40 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{pmatrix} \right]$. Check if X_1 and $X_1 + 3X_2 - 2X_3$ are independent and also find its expectation.

   6+4

5. (a) Show that in multiple regression theory, the residual variance $s^2_{1.23\cdots p}$ is decreasing sequence in p.

(b) Deduce the moment generating function of Bivariate normal (4, 6, 9, 2, 0.7).

5+5

6. Write short note on any **two** of the following.

   (a) Multiple correlation coefficient
   (b) Mann Whitney U test
   (c) Multivariate moment generating function

5+5

B.Sc. (Honours) Examination, 2022
Semester-VI
Statistics
Course: CC-14B
Multivariate Analysis & Nonparametric Methods

Time: 2 hrs Full Marks:20

Answer **all** the questions. Tables are attached at the bottom of the questions.

1. If $r_{12} = .80$, $r_{13} = -.40$ and $r_{23} = -.56$, find the values of $r_{2.13}$, $r_{21.3}$ and $s_{1.23}$. Can you say what amount of total variance of dependent variable is explained by x_1 and x_3?

   4

2. Cholesterol measurements were taken on 20 heart-attack patients on their cholesterol levels. For each patient, measurements were taken 0, 2, and 4 days following the attack. Treatment was given to reduce cholesterol level. Let X_1 =0-day chol.level, X_2 =2-day chol.level, X_3 =4-day chol.level be the variables for which the mean vector is $(259.5, 230.8, 221.5)'$ and the covariance matrix is $\begin{pmatrix} 2276 & 1508 & 813 \\ 1508 & 2206 & 1349 \\ 813 & 1349 & 1865 \end{pmatrix}$.

   (a) Find the mean and variance of $X_2 - X_3 + 2X_1$.

   (b) Assume that the underlying distribution from which each observation came is multivariate normal with the mean vector and covariance matrix already given above. Predict the cholesterol level of 4-day when 0-day cholesterol level is 230 and 2-day cholesterol level is 210.

   (c) Find the probability that 4-day level cholesterol will be more than 240.

   3+3+2

3. Ten 4-year-old boys and ten 4-year-old-girls were observed during two 15 minute play sessions and each child's play during these two periods was scored as follows for incidence and degree of aggression.

Boys:	86	69	72	65	113	65	118	45	141	104
Girls:	55	40	22	58	16	26	120	15	104	66

   Test the hypothesis that there were gender differences in the amount of aggression shown.

   4

4. We observe the following 8 data points:
   1.41 0.26 1.97 0.33 0.55 0.77 1.46 1.18
   Is there any evidence to suggest that the data were not randomly sampled from a Uniform(0,2) distribution?

   4

B.Sc. (Honours) Examination, 2022
Semester-VI
Statistics
Course: DSE-3
(Operations Research [Theory])
Time: 3 Hours **Full Marks: 40**

Questions are of value as indicated in the margin
Notations have their usual meanings

$\underline{Group - A}$ $5 \times 2 = 10$

1. Answer any five of the following questions with proper justification.

   (a) Discuss convex set by citing an example.

   (b) What do you mean by a feasible solution?

   (c) Write down a linear programming problem with three variables and four constraints. Hence convert the primal into its dual problem.

   (d) How you can define net evaluations in the context of transportation problem?

   (e) Under which condition(s) an assignment problem has degenarate solution.

   (f) What is a pay-off matrix in game theory?

   (g) What are the various types of costs associated with inventory control models?

   (h) How to find Lorentz curve in an ABC analysis?

$\underline{Group - B\ (Answer\ any\ three\ questions)}$ $3 \times 10 = 30$

2. (a) Briefly write down various types of operations research problem.
   (b) Solve the linear programming problem (if possible) with dual simplex method

$$Maximize,\ z = 6x_1 + 7x_2 + 3x_3 + 5x_4$$

Subject to,

$$5x_1 + 6x_2 - 3x_3 + 4x_4 \geq 12$$
$$x_2 + 5x_3 - 6x_4 \geq 10$$
$$2x_1 + 5x_2 + x_3 + x_4 \geq 8$$
$$x_1, x_2, x_3, x_4 \geq 0..$$ 4+6

3. (a) Discuss the mathematical formulation of a transportation problem by citing an example.
   (b) Obtain the IBFS to the following transportation problem by matrix (cost) minima method then find out an optimal solution and corresponding cost of the transporta tion. 5+5

$$\begin{bmatrix} & D_1 & D_2 & D_3 & D_4 & a_i \\ O_1 & 5 & 4 & 6 & 14 & 15 \\ O_2 & 2 & 9 & 8 & 6 & 4 \\ O_3 & 6 & 11 & 7 & 13 & 8 \\ b_j & 9 & 7 & 5 & 6 & \end{bmatrix}$$

4. (a) For an assignment problem, suppose (c_{ij}) is a cost matrix of order n. Now if α_i, β_j are arbitrary constants, then the matrix $(c^*_{ij}) = (c_{ij} - \alpha_i - \beta_j)$ and (c_{ij}) have identical solutions.

   (b) Briefly write down the steps of Hungarian method for solving an assignment problem. Hence solve the following assignment problem and find the optimal assignment profit from the profit matrix below. 4+(1+5)

$$\begin{bmatrix} & A & B & C & D & E \\ 1 & 5 & 7 & 6 & 8 & 7 \\ 2 & 10 & 7 & 9 & 11 & 7 \\ 3 & 5 & 12 & 11 & 13 & 7 \\ 4 & 11 & 9 & 15 & 10 & 7 \\ 5 & 5 & 11 & 8 & 11 & 11 \end{bmatrix}$$

5. (a) What is the role of post optimality analysis in linear programming problem? In standard notation, suppose we change a component of the requirement vector b. Discuss the optimality of new basic feasible solution and the change in optimal cost.

   (b) Solve the following 2×4 game graphically. Here A is the row player having strategies A_1, A_2 and B is the column player having strategies B_1, B_2, B_3, B_4.

$$\begin{bmatrix} & B_1 & B_2 & B_3 & B_4 \\ A_1 & 3 & 2 & -1 & 4 \\ A_2 & 2 & 5 & 6 & -2 \end{bmatrix}$$

4+(2+4)

6. (a) Why it is essential for mantaining an inventory for almost all the companies?

   (b) Write down five limitations of ABC Analysis. 5+5

B.Sc. (Honours) Examination, 2022
Semester-VI
Statistics
Course: DSE-3B
(Practical on Operations Research)

Time: 2 Hours **Full Marks: 20**

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **all** questions

1. Solve the LPP by algebraic method

$$Maximize,\ z = 5x_1 + x_2 + 3x_3$$

Subject to,

$$x_1 + 2x_2 - 2x_3 \leq 30$$

$$x_1 + 3x_2 + x_3 \leq 36$$

$$x_1, x_2, x_3 \geq 0.$$ 6

2. Solve graphically the following LPP

$$Minimize,\ z = -2x_1 + x_2$$

Subject to,

$$x_1 + x_2 \geq 6$$

$$3x_1 + 2x_2 \geq 16$$

$$x_2 \leq 9$$

$$x_1, x_2 \geq 0.$$ 4

3. Four products are produced in three machines and their margins are given by the table below with capacity a_i and requirement b_j. Find out an optimal allocation so that the profit is maximized. Also discuss about its multiple optimal solutions. 7

$$\begin{bmatrix} & P_1 & P_2 & P_3 & P_4 & a_i \\ M_1 & 6 & 4 & 1 & 5 & 14 \\ M_2 & 8 & 9 & 2 & 7 & 18 \\ M_3 & 4 & 3 & 6 & 2 & 7 \\ b_j & 6 & 10 & 15 & 8 & \end{bmatrix}$$

4. Practical note book and viva-voce. 3