Birla Institute of Technology and Science, Pilani

Statistics and Basic Econometrics, **Code: MPBAG505**

First Semester 2023-2024, Comprehensive Examination, IC: Dr. Achint Nigam

==========

Instructions

- Maximum time = 180 minutes; Maximum marks = 70. Marks are mentioned against the questions.
- Attempt all questions. This paper has 2 pages and 8 questions.
- If needed, you can make any suitable assumptions you may like. Please mention them.
- Give tabulated responses wherever possible.
- Closed book examination. A formula sheet is given.
- Tables can be requested from the invigilator.

==========

1. The following data show the rankings of 11 states based on expenditure per student (ranked 1 highest to 11 lowest) and student–teacher ratio (ranked 1 lowest to 11 highest).

State	Expenditure per student	Student-Teacher Ratio
Delhi	9	10
UP	5	8
MP	4	6
Raj.	2	11
Kerala	6	4
TN	11	3
Telangana	1	1
Punjab	7	2
Haryana	8	7
J&K	10	5
WB	3	9

   a. What is the rank correlation between expenditure per student and student–teacher ratio? [5]
   b. State your hypotheses and at the α = .05 level, does there appear to be a relationship between expenditure per student and student-teacher ratio? [5]

2. Consider the following data for two variables, x, and y (dependent)

x	2	3	4	5	7	7	7	8	9
y	4	5	4	6	4	6	9	5	11

   a. Develop the estimated regression equation relating *x* and *y*. [2]
   b. Show in a tabular form $\hat{y}$, residuals and the standardized residuals. [8]
   c. Using the graph paper, plot the **standardized residuals** versus **x** for the estimated regression equation developed in part (a). Does the model assumption of constant variance appear to be satisfied? [5]

3. In a regression analysis involving 27 observations, the following estimated regression equation was developed: $\hat{y} = 25.2 + 5.5x_1$, SST = 1550 and SSE = 520
   a. At α = .05 level, test whether x_1 is significant. **Clearly mention *df* and critical value.** [5]

   Suppose that variables x_2 *and* x_3 are added to the model and the following regression equation is obtained $\hat{y} = 16.3 + 2.3x_1 + 12.1x_2 - 5.8x_3$, SST = 1550, SSE = 100
   b. At α = .05 level, determine whether x_2 and x_3 contribute significantly to the model. **Clearly mention *df* and critical value.** [5]

Birla Institute of Technology and Science, Pilani

Statistics and Basic Econometrics, **Code: MPBAG505**

First Semester 2023-2024, Comprehensive Examination, IC: Dr. Achint Nigam

4. In the table below you are given marks of 5 students each in three sections of a statistics course.

S. No.	Section 1	Section 2	Section 3
1	25	9	23
2	35	17	13
3	31	28	33
4	24	36	15
5	14	32	19

   a. You need to find out if the marks scored by students in these three sections are different on average or not. Clearly show your calculations in tabular form. Use α = .05 level. [7]
   b. Discuss how the same problem can be solved using linear regression analysis. *There is no need for calculation for this part. Just show the **complete** input datatable for regression.* [3]

5. Using the data from Q2, assume there is a third variable z, whose data is given below. Compute sr_x. Clearly show all your steps. [10]

z	4	3	4	3	5	5	6	6	7

6. What do you understand by multiple comparison procedures? Give two examples. *The formula is not needed.* [5]
7. Conceptually what is the difference between correlation, semi-partial correlation, and partial correlation? [5]

8. Differentiate between forward selection and backward elimination variable selection procedures [5]