# Assignment MST-017

for

# M.Sc. (Applied Statistics) (MSCAST)

Valid from January 2026 to December 2026

**SCHOOL OF SCIENCES**
Indira Gandhi National Open University
New Delhi - 110068

**Dear Learner,**

Welcome to the M.Sc. (Applied Statistics) Programme.

As per the university guidelines, you need to complete the assignment for each theory course. Note that there are no assignments for lab courses in the MSCAST programme, namely, MSTL-011, MSTL-012, MSTL-013, MSTL-014, and MSTL-015. You should remember that writing answers to an assignment's questions will improve your writing skills and prepare you for the term-end examination.

**It is compulsory to submit the assignments within the stipulated time to be eligible to appear in the term-end examination.** You will not be allowed to appear for the term-end examination for a course if you do not submit the assignment for that course by the due date. As per the University guidelines, if you appear in the term-end examination of a course without submitting its assignment, the result of the term-end examination is liable to be cancelled/ withheld.

**The assignments constitute the continuous component of the evaluation process and have 30% weightage in the final grading.**

Before you write the assignments, you are advised to first go through the self-learning material for that course and then prepare the assignments carefully by following the instructions pertaining to the assignments. Your responses should not be a verbatim reproduction of the textual materials provided for self-learning purposes, but it should be in your own words.

If you have any doubts or problems pertaining to the course material and assignments, contact the programme in charge or the academic counsellor at your study centre. If you still have problems related to this assignment, feel free to contact the course coordinator.

Wishing you all the best in successfully completing the programme.

**(Dr. Neha Garg)**
**Course Coordinator, MST-017**
**Email:** nehagarg@ignou.ac.in

## Instructions:

- Submit the assignments within the stipulated time. Otherwise, you will not be permitted to appear for the term-end examination.
- Solve the latest assignments uploaded for the current year/session.
- Read the instructions related to the assignments mentioned in the Programme Guide.
- Use only A-4 size paper to write your responses. It is mandatory to write all assignments neatly in your own handwriting. Typed or printed copies of the assignments will not be accepted. Note that you may use the printout only if a question specifically asks for the output of a program in MST-015 and MST-024.
- All questions given in the assignments are compulsory for each course.
- Express your response in your own words. You are advised to restrict your response based on the marks assigned to it. This will also help you to distribute your time in writing or completing your assignments on time.
- Securely fasten multiple pages together (you can staple or tie them) and number them carefully for each assignment separately.
- Do not forget to enclose the assignment question sheet of that course after the cover page of the assignment response (answer sheets). It is not compulsory to write each question separately before answering the question. Mention the question number for each answer.
- The solved assignment must be submitted at the Study Centre allotted to you before the due date set by the University. Please check the IGNOU website for updated information regarding the due date of assignment submission.
- You are advised to mention all information on the first page of the assignment response sheet, given on the next page.
- **Keep a copy of the assignment answer sheets with you before submission for future reference.**

# ASSIGNMENT CODE: MST-017/TMA/2026

**NAME:** _____

**ENROLLMENT NO:** _____

**ADMISSION CYCLE:** _____

**PROGRAMME CODE:** MSCAST

**COURSE CODE:** MST-017

**COURSE TITLE:** APPLIED REGRESSION ANALYSIS

**REGIONAL CENTRE CODE:** _____

**STUDY CENTRE CODE:** _____

**ADDRESS:** _____

_____

_____

**CONTACT NUMBER:** _____

**EMAIL ID:** _____

**DATE OF SUBMISSION:** _____


**School of Sciences**

**Indira Gandhi National Open University**

**Maidan Garhi, New Delhi-110068 (INDIA)**

**TUTOR MARKED ASSIGNMENT**

# MST-017: Applied Regression Analysis

**Course Code: MST-017**

**Assignment Code: MST-017/TMA/2026**

**Maximum Marks: 100**

**Note: All questions are compulsory. Answer in your own words.**

**1.** State whether the following statements are true or false and also give the reason in support of your answer. **(5×2=10)**

**(i)** The model $Y = \beta_0 X^{\beta_1} + \varepsilon$ is a linear model.

**(ii)** If total and residual sum of squares for a dataset are 124 and 15.5, respectively, then the coefficient of determination is 0.875.

**(iii)** We use Cook's distances to identify the autocorrelation of the error terms.

**(iv)** If the coefficient of determination for the auxiliary regression model of the explanatory variable $X_1$ is 0.8, the VIF for $X_1$ will be 5.

**(v)** We use baseline category logit models when we have ordered categories of a response variable.

**2.** To investigate the effect of Age (in years) and Smoking Status (Group 1: Non-Smokers, Group 2: Smokers) on Systolic Blood Pressure (in mmHg), the following data were collected:

| Non-smoker | | | Smoker | | |
|---|---|---|---|---|---|
| **Patient** | **Age (X)** | **Systolic Blood Pressure (Y)** | **Patient** | **Age (X)** | **Systolic Blood Pressure (Y)** |
| 1 | 40 | 120 | 1 | 35 | 128 |
| 2 | 45 | 124 | 2 | 42 | 135 |
| 3 | 50 | 130 | 3 | 48 | 142 |
| 4 | 55 | 132 | 4 | 53 | 148 |
| 5 | 60 | 138 | 5 | 58 | 155 |
| 6 | 65 | 142 | 6 | 62 | 158 |
| 7 | 70 | 145 | 7 | 68 | 165 |
| 8 | 75 | 150 | | | |

**(i)** Fit a multiple regression model.

**(ii)** Obtain the individual regression models for Non-Smokers and Smokers separately based on the main model.

**(iii)** Test the significance of the overall fitted regression model at the 5% level of significance.

**(8+4+8)**

**3.** Fit a simple linear regression model on the data given in Question **2** for non-smokers only and obtain the followings:

**(i)** Diagonal of the hat matrix. Also, check the leverage points, if any.

**(ii)** Cook's Distances. Also, verify the influence points, if any.

**(6+14)**

**4** A logistics company wants to predict the delivery efficiency score of its drivers based on two primary factors: years of driving experience and average hours of sleep per night. The data for 10 drivers are recorded as follows:

| Driver | Experience ($X_1$) | Sleep ($X_2$) | Efficiency (Y) |
|---|---|---|---|
| 1 | 2 | 8 | 75 |
| 2 | 5 | 7 | 82 |
| 3 | 10 | 6 | 88 |
| 4 | 1 | 5 | 60 |
| 5 | 8 | 8 | 95 |
| 6 | 4 | 7 | 78 |
| 7 | 3 | 6 | 72 |
| 8 | 12 | 7 | 98 |
| 9 | 7 | 4 | 80 |
| 10 | 6 | 9 | 90 |

Evaluate all possible regression models using adjusted $R^2$ and AIC. Also, obtain best subset model.

**(25)**

**5.** The following table shows the concentration of a new antibiotic, the total number of infected subjects, and the number of subjects who showed complete recovery:

| S. No. | Concentration ($x_i$) | Recovered ($y_i$) | Total Subjects ($n_i$) |
|---|---|---|---|
| 1 | 2 | 10 | 50 |
| 2 | 4 | 15 | 45 |
| 3 | 6 | 22 | 40 |
| 4 | 8 | 35 | 60 |
| 5 | 10 | 68 | 80 |

**(i)** Fit a logistic regression model considering $\hat{\beta}_0^0 = -2.148$ and $\hat{\beta}_1^0 = 0.36$, as the initial values of the parameters. Perform calculations for only one iteration using the Newton-Raphson method.

**(ii)** Test the significance of the fitted model using the Hosmer-Lemeshow test at 5% level of significance.

**(15+10)**