Assignment 8
Due on 2020-11-11, 23:59 IST

Instructions for Q1 to Q10:
This is the final exam on a course that contains the data of students. This data is artificially coded data for this course and does not belong to any actual study.

1. Mid_vent: Percentage - This represents the marks obtained in mid-semester examination. (Type: Integer)
2. Attendance_percentage: The student's attendance in the online interaction.
3. Mid_sem_marks: This represents the marks obtained in mid-semester examination. (Type: Integer)
4. End_sem_marks: This is the final marks obtained by the student.

Use this data and perform linear regression (multiple linear regression) on the data to answer the following questions. Kindly take the exact values of the parameters into consideration.

(Round your answers correct to two decimal places)

1. What is the value of intercept?
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 0.53, 0.58

2. What is the value of the coefficient of "Mid_sem_marks"?
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 0.6, 0.62

3. The Precision of the model is _______.
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 0.85, 0.82

4. What is the value of coefficient of "End_sem_marks"?
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 0.62, 0.61

5. A researcher has developed a new ITS, which teaches calculus to 12th students. To check the effectiveness of this intervention, he took a random sample of 12 students.
   - Data collected from the students using the ITS is given below only.
   - Data collected from previous studies among 10 students.
   - Data collected from five previous studies belonging to the same class of calculus.
   - Data collected from the ITS is used because both are science subjects.
   - No, the person is incorrect.
   - Accepted/Answers: Data collected from previous studies belonging to the same class of calculus.

Use the given table to answer Q11 to Q15

<table>
<thead>
<tr>
<th>ID</th>
<th>Mid_sem_marks</th>
<th>End_sem_marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>65</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>80</td>
</tr>
<tr>
<td>7</td>
<td>43</td>
<td>75</td>
</tr>
<tr>
<td>8</td>
<td>33</td>
<td>67</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>47</td>
<td>77</td>
</tr>
</tbody>
</table>

Develop a linear regression model between end_sem_marks and Mid_sem_marks using true data ID 1 to 5 and answer Q11, Q12, and Q13.

(Round your answers correct to two decimal places)

11. What is the value of intercept?
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 0.5, 0.53

12. What is the value of slope?
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 1.5, 1.56

13. For the above problem, calculate the end_sem_marks for individual ID 10.
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 15.35, 15.03

14. In the analysis of logistic regression, ignored function are given as following:
   \[ f(x) = \frac{1}{1 + e^{-x}} \]
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 1.5, 1.5

15. What does parameter b represent?
   - No, the person is incorrect.
   - Accepted/Answers: Type: Target 0.5, 0.5

16. Which of the following are true for logistic regression?
   - It is a classification algorithm
   - It is a regression algorithm
   - Dependent variable should be categorical while independent variable should be numerical
   - Independent variable should be categorical while dependent variable should be numerical
   - No, the person is incorrect.
   - Accepted/Answers: Dependent variable should be categorical while independent variable should be numerical.