The Open University of Sri Lanka
B.Sc/B.Ed. DEGREE, CONTINUING EDUCATION PROGRAMME
No Book Test (NBT) 2017/2018
Level 05 - Applied Mathematics
ADU5301– Regression Analysis I

Date: 27.01.2019 **4.15pm to 5.15pm**

Instructions

- This examination is of **one hour** duration.
- Answer **all** questions.
- Each of the two questions is allocated fifty marks.
- Non programmable calculators are permitted.

1. The following summary statistics were computed from the dried weights of 25 plants (in mg) at the age of six months, y where x denote the amount of fertilizer added (in mg).

   $\sum x_i = 125.0, \sum y_i = 233.7, \sum x_i y_i = 1548.3, \sum x_i^2 = 989.0, \sum y_i^2 = 2623.0.$

   The amounts of fertilizer added to the samples had varied from 0mg to 12mg. A simple linear regression model is to be fitted to the data using the method of least squares, with the amount of fertilizer as the predictor variable.

   i) Obtain least squares estimates for the slope and intercept of the population regression line.

   ii) Write down the equation of the fitted line.

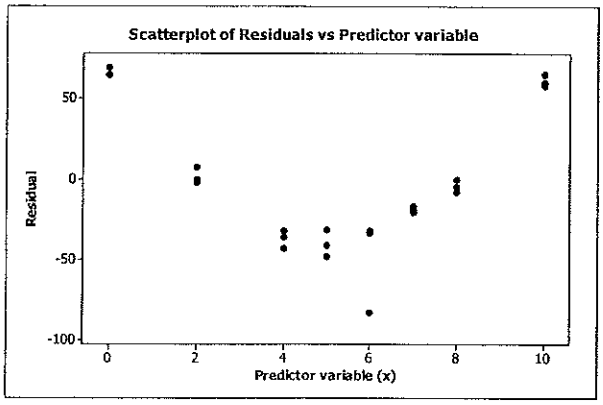
   iii) Estimate the expected dried weight of a plant that had received *4*mg of the catalyst.

   iv) State whether each of the following statements is true or false. If you decide that the statement is false, give reasons.

      a) The fitted line obtained from the method of least squares passes through more points in a scatter plot of data compared to any other fitted line.

      b) An observation that lies on the fitted line does not have any random error.

2. A researcher measured the reaction time, y on 25 chemical samples with different amounts of catalyst, x added. The amounts of catalyst used for the study had varied from 0mg to 10mg per sample. A simple linear regression model was fitted to the data using the method of least squares. Plot of residuals obtained from this model fit is given below.



Using the information conveyed by the residual plot, state whether you agree with each of the following statements or not. If you do not agree, give reasons.

i) The regression function used by the researcher does not provide a satisfactory fit to the data.

ii) The constant variance assumption made on the random errors is not appropriate.

iii) The researcher has used at most three observations at each level of the predictor variable.

iv) The model fits best for observations collected on samples with 5mg of the catalyst added.

v) The researcher should have used the amount of catalyst as the predictor variable.