The Open University of Sri Lanka

B.Sc/B.Ed. DEGREE, CONTINUING EDUCATION PROGRAMME

No Book Test (NBT) 2021/2022

Level 05 - Applied Mathematics

ADU5301– Regression Analysis I

Date: 12.02.2022 **9.00 am to 10.00 am**

Instructions

- This examination is of **one-hour** duration.
- Answer **all** questions.
- Each of the two questions is allocated equal marks, distributed as indicated. The total marks allocated is 200 and the earned marks will be converted to a mark out of 100.
- Non programmable calculators are permitted.

1. In a study to assess the effect of temperature (x), on the rate of a chemical reaction (y), a researcher measured the reaction rates on 35 samples. The temperatures used for data collection were 20 ^0C, 25 ^0C, 30 ^0C, 35 ^0C, 40 ^0C, 45 ^0C and 50 ^0C. The researcher had used 5 replicates at each temperature level. The summary statistics calculated are given below.

$$\sum x_i = 1225\ ;\ \sum y_i = 267.4\ ;\ \sum x_i^2 = 46375;\ \sum y_i^2 = 2116.76;\ \sum x_i y_i = 9843.0;$$

The researcher wants to fit a simple linear regression model to the yield using the method of least squares, with temperature as the predictor variable.

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

(30 marks)

ii) Write down the equation of the fitted line, based on the method of least squares.

(10 marks)

iii) Define the term replicate, in relation to this study and describe two advantages of having replicates.

(25 marks)

iv) Using the fitted model, estimate the reaction rate of a chemical sample, if the sample is maintained at 32 ^0C.

(10 marks)

v) Suppose the researcher had noted an unusually low reaction rate in one of the samples that was maintained at $35\ ^0C$. A student stated that if the data collected on this sample were not used estimate for the slope would be very different to the value obtained in part (i), Do you agree with this statement? Give reasons for your answer.

(25 marks)

2. A student fitted the following two models for a set of data with 40 observations, using the method of least squares and calculated the residuals from each model fit.

Model 1: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$

Model 2: $y = \beta_0 + \beta_1 x^2 + \epsilon$

i) Suppose from a residual analysis, a student concluded that Model1 fits well to the data whereas Model2 does not adequately describe the data. Clearly describe the residual analysis that would allow the student to make this conclusion.

(25 marks)

ii) Suppose from a residual analysis, a student found that both models satisfy the model assumptions.

a) Clearly describe the assumptions the student has to make to fit Model2.

(25 marks)

b) Which model would you recommend to use? Give reasons for your answer.

(25 marks)

iii) The student had found that the estimates for β_0 and β_1 are 12.1 and 0.89 respectively. If the values of the response and predictor variables for a data point are 2.9 and 3 respectively, calculate the residual for this data value.

(10 marks)

iv) Based on the value calculated in part (iii), state with reasons whether the data point agrees with the chosen model fit or not.

(15 marks)