THE OPEN UNIVERSITY OF SRI LANKA

B.Sc. /B.Ed. Degree Programme, Continuing Education Programme

APPLIED MATHEMATICS - LEVEL 05

ADU5301– Regression Analysis I

Open Book Test (OBT) 2021/2022

**Date: - 23.12.2022** **Time: 9.00 a.m. to 10.00 a.m.**

**Instructions**

- This examination is of **one hour** duration.
- Answer **all** questions.
- Each of the two questions is allocated fifty marks. **In the first question, each part is given 25 marks equally distributed among the subparts. In the second question, marks are allocated as indicated.**

1) A researcher measured the weight (mg) of the precipitate obtained by adding known amounts of a certain chemical compound to 40 samples. Suppose the researcher wants your advice to model the relationship between the weight of the precipitate and the amount of chemical compound added.

   a) Which variable would you choose as the explanatory variable? Give reasons for your choice.

   (5 marks)

   b) Briefly explain what is meant by the random error in an observation collected in this study.

   (5 marks)

   c) On theoretical grounds, the researcher expects that the regression function $f(x, \underline{\beta}) = \beta_0 + \beta_1\sqrt{x}$ is appropriate for the mean response.

      i) Is the regression function chosen by the researcher, suitable for a simple linear regression model? Justify your answer.

      (5 marks)

      ii) What preliminary analysis would you recommend, if the model is to be fitted using the method of least squares? Clearly explain what you plan to identify from the preliminary analysis you recommended.

      (10 marks)

2) The following summary statistics were obtained from the data collected on the length of training, $x$ (weeks), and the raw material wastage, $y$ (grams) of 40 workers.

$$\sum x_i = 258;\ \sum y_i = 248.2;\ \sum x_i^2 = 2212;\ \sum y_i^2 = 2508.2;\ \sum x_i y_i = 1542.6$$

a) Calculate the Pearson correlation coefficient.

(15 marks)

b) What do you conclude from the value calculated in part (a)? Clearly describe any assumptions that you use to arrive at the stated conclusions.

(05 marks)

c) Let the value of the Pearson correlation coefficient calculated in part (a) be $r$. If the researcher later noticed that the scale used to measure the response had underestimated all $y$ values by 0.02 grams. A student claimed that the Pearson correlation coefficient between the variables after correcting for the error in underestimation will be $r+0.02$. Do you agree with this statement? Give reasons for your answer.

(05 marks)