The Open University of Sri Lanka

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

Open Book Test (OBT) 2024/2025

Level 05 - Applied Mathematics

ADU5301– Regression Analysis I

Date: February 08, 2025 **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 equal marks, distributed as indicated.
- Non programmable calculators are permitted.

1. A study assessed reading and writing skills in 200 children before and after a voluntary training program, which lasted up to 30 days. After completing the training, each child received marks for both reading and writing skills. Scatter plots of marks (y) against the length of training (x) indicated that a regression function of the form $f(x, \underline{\beta}) = \beta_0 + \beta_1 x$, where $\underline{\beta} = (\beta_0, \beta_1)^T$ is appropriate for the marks in each skill. The slope of the line for reading skills was approximately twice as the slope for the writing skills.
Based on these findings, a student made the following conclusions. State whether you agree or disagree with each conclusion and provide justification for your answer:
   a) If a student increases the length of training by two days, the expected increase in reading test marks will be approximately twice as large as the increase in writing test marks.
   b) Improvements in reading skills will cause improvements in writing skills.
   c) A simple linear regression model with reading skills as the predictor variable and writing skills as the response variable can provide good estimates of writing skill marks based on reading skill marks.
   d) The available information is insufficient to determine whether the Pearson correlation between reading and writing skills will be positive or negative.

e) Marks for writing skills of a student can also be predicted using a simple linear regression model fitted to marks for reading skills with the length of training as the predictor variable.

(5 × 10 = 50 marks)

2) i) Let x denote a predictor variable and β_0 and β_1 denote unknown parameters. State whether each of the following regression functions is suitable for a simple linear regression model fit or not. In each case, give reasons for your answer.

a) $f(x, \beta_1) = \beta_1 x$

b) $f(x, \underline{\beta}) = \beta_0 + \beta_0 \beta_1 x$, where $\underline{\beta} = (\beta_0, \beta_1)^T$

c) $f(x, \underline{\beta}) = \beta_0 x + \beta_1 x^2$, where $\underline{\beta} = (\beta_0, \beta_1)^T$

(3 × 5 = 15 marks)

ii) A researcher recorded the preservation times (in days) of 30 samples of a certain food product after adding known amounts of a food preservative. The amounts of preservative used for the study had varied from 1 gram to 5 grams per 400 grams of the food product. Suppose the researcher seeks your advice to fit a simple linear regression model, including an intercept term, for the response variable y using x as the only predictor variable.

a) Which variable would you choose as the predictor variable? Give reasons for your answer.

(5 marks)

b) Write down the model equation for the response y in the usual notation.

(5 marks)

c) The following summary statistics were computed from the data collected in the study.

$n = 30, \sum x = 90; \sum x^2 = 330; \sum y = 170.9; \sum y^2 = 1061.21; \sum xy = 585.2$

Based on the model equation given in part (b), estimate the mean change in the response associated with one unit of change in the predictor variable.

(25 marks)