The Open University of Sri Lanka

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

Open Book Test (OBT) 2016/2017

Level 03 - Applied Mathematics

APU2141/APE4141 – Regression Analysis I

Date: - 07th October 2017   **10.30am to 11.30am**

Instructions

- This examination is of **one hour** duration.
- Answer **all** questions.
- Each of the two questions is allocated fifty marks.

1. State whether each of the following statements is true or false. If the statement is false, explain why it is false.

   a) Pearson correlation coefficient between two variables is almost zero implies that a simple linear regression model cannot adequately describe the data.

   b) If the Pearson correlation coefficient between two variables is positive, the estimate for the slope parameter based on the same data should also be positive.

   c) In a scatter plot of the data, if the points are tightly clustered about a line, the Pearson correlation coefficient should be close to ± 1.

   d) In the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$, the parameter β_1 represents the change in the response associated with one unit of change in the explanatory variable.

   e) In a simple linear regression model, random errors in observations that lie above the fitted line are positive.

2. State whether the research objective in each of the following can be achieved by fitting a simple linear regression model. If the objectives can be achieved, state the variable you choose as the explanatory variable X, in the usual notation. If not, give reasons. Where applicable, you may assume that the assumptions on the error term are reasonable.

a) The mean yield of a chemical reaction is known to be related to the temperature t (^0C) according to the equation $E(y) = \alpha_0 + \alpha_1(t-20)^2$, where α_0, α_1 are some unknown parameters. The researcher wants to estimate the parameters based on the yields measured on 50 samples at temperatures (^0C) 0, 5, 10, 15, 20, 25, 30, 35, 40 and 45 with 5 replicates at each temperature.

b) A researcher wants to find out whether the regression function $\mu = \gamma_0 + \gamma_1\sqrt{age}$ can give a valid estimate of the height (cm) of a medicinal plant at the age of six months, based on the heights measured on 40 plants at ages 3, 5, 8, 10 and 12 months where heights of eight plants are measured at each age; here μ denotes the mean response.

c) A researcher wants to find out whether the function $\mu = \alpha_0 + e^{\alpha_1 t}$ adequately describes the expected yield μ, of a chemical reaction at temperature t based on 80 observations collected at known temperatures.

d) A researcher wants to find out whether the expected production cost (P) is related to the cost of raw material (R) and the cost of labour (L) according to the equation $P = \alpha_0 + \alpha_1 R + \alpha_2 L$, where α_0, α_1 and α_2 are parameters to be estimated based on the past data.

e) Based on the times taken by 200 persons who had undergone a certain training programme for different lengths of time, a researcher wants to find out whether the function $\mu = \dfrac{\alpha_0 + \alpha_1 t}{\alpha_0 t}$ is appropriate to describe the mean time to complete the task μ, where t denotes the length of training and α_0 and α_1 denote parameters to be estimated.