00207

THE OPEN UNIVERSITY OF SRI LANKA

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

APPLIED MATHEMATICS-LEVEL 04

APU2141/APE4141- REGRESSION ANALYSIS I

FINAL EXAMINATION 2014/2015

Duration: Two Hours.

Date: 21.10.2015	**Time: 9.30p.m- 11.30p.m**

Answer FOUR questions only.

Instructions:

- **This question paper consists of 06 questions. Answer only four questions.**
- **Statistical Tables are provided. When reading values, you may use the closest degrees of freedom given in the table.**
- **Non-programmable calculators are permitted.**

1. A researcher interested in studying the effectiveness of a catalyst on the reaction time of a chemical process, added different known amounts (mg) of the catalyst and measured the reaction times (minutes). The data collected are given in the accompanying table.

Amount of catalyst (x)	0	0	0	1	1	1	2	2	2	3	3	3
Reaction time (y)	1.0	1.5	1.5	1.5	2.0	2.0	2.5	2.5	3.0	4.5	5.0	5.5

\The equation of the fitted simple linear regression model for this data, using the method of least squats is $y = 0.933 + 1.183x$.

i) Compute the fitted values.

ii) Compute the residuals.

iii) Estimate the random variation in the reaction time relying on the fitted model.

iv) Construct a suitable graphical summary that can be used to examine whether the chosen regression function is appropriate.

v) Clearly state all the findings from the graphical summary constructed in part (iv).

2. The following descriptive statistics were computed from the lengths (cm) and weights (grams) measured on 20 Sri Lanka blue magpies.

`Variable	Sample mean	Standard deviation
Weight	186.37	5.26
Length	42.0	3.81

Pearson correlation coefficient between weight and height = 0.93.

A simple linear regression model is fitted, using the method of least squares, with the weight and height as the response and predictor variables respectively. An analysis of variance (ANOVA) table is constructed for the model fit,

i) Clearly describe what one can conclude from the value of the Pearson correlation coefficient,

ii) Estimate the slope parameter.

iii) Find the regression sum of squares.

iv) Find the total sum of squares.

v) Estimate the percentage of variance in the weights that can be explained using length.

3. A researcher interested in studying the effect of a chemical compound added to diesel on the mileage travelled, recorded the amount of compound added per litre of diesel (mg) and the mileage per litre (km), y, The amounts of compound used per litre (x) were 0mg, 2mg, 4mg, 6mg and 10mg. Three replicates were collected at each level of the predictor variable.

The researcher fitted a simple linear regression model using the method of least squares and obtained the following results.

fitted equation : $y = 10.671 + 0.529x$

mean squared error = 0.788

i) Estimate the standard error in the estimate for the slope parameter.

ii) Construct a 95% confidence interval for the slope parameter.

iii) Using level of significance as 0.05, test the validity of the claim that the added compound has no significant effect and clearly state the findings.

iv) Clearly state the assumptions one has to make to test the validity of the claim stated in part (iii).

4. The following summary statistics were computed from the data collected in a study to assess the effectiveness of a training programme. The response and the predictor variables are the time to accomplish a specific task (minutes) and the length of training (days). Among the 20 participants of the study, three persons had no training. The longest training was for seven days.

   The summary statistics computed from the data are given below.

   $\sum x = 59, \sum y = 91.0, \sum x^2 = 263, \sum y^2 = 532.0, \sum xy = 170.0.$

   The researcher fitted a simple linear regression model with length of training as the predictor variable.

   i) Find the least squares estimates for the slope and intercept parameters.

   ii) Write down the equation of the fitted model.

   iii) If a worker is given five days of training, estimate the expected time he will take to accomplish the task.

   iv) Based on the fitted model, a student concluded that a worker with two weeks of training will accomplish the task almost fifteen ~~hours~~ minutes earlier compared to a worker who had not undergone the said training. Comment on the conclusion.

5. A researcher recorded the dissolving time (in minutes) of a compound in an acidic medium on 28 samples kept at different temperatures (^0C). The following summary statistics were computed from the data.

   $\sum x = 603, \sum y = 441.0, \sum x^2 = 14055, \sum y^2 = 8185.0, \sum xy = 10627.0.$

   The researcher wants to test whether the temperature has a significant effect on the dissolving time or not

   i) Write down the null and the alternative hypotheses you would test to address the researcher's objectives.

   ii) Compute the least squares estimate for the slope parameter.

   iii) Construct an analysis of variance (ANOVA) table that can be used to address the researcher's objectives.

   iv) Using the ANOVA table constructed in part (iii), test the hypothesis using a 0.05 significance level, and clearly state the findings.

6. A part of the analysis of variance (ANOVA) table constructed by fitting a simple linear regression model using the method of least squares for the data collected on the dried weight (mg) of a medicinal plant using the period of drying in the sunlight (minutes) as a predictor variable is given below.

Source of variation	Sum of squares	Degrees of freedom
Regression	16.654	(b)
Residual	(a)	(c)
Total	18.293	21

i) What is the sample size used for the study?

ii) Find the values indicated by (a), (b) and (c).

iii) Estimate the random variation in the response variable.

iv) Compute the coefficient of determination, R^2, and clearly explain what it measures in relation to this study.

v) Assuming that there is a negative linear association between the response and the predictor variables, find the Pearson correlation coefficient,