



(Knowledge for Development)

**KIBABII UNIVERSITY**  
**UNIVERSITY EXAMINATIONS**  
**2016/2017 ACADEMIC YEAR**  
**THIRD YEAR FIRST SEMESTER**  
**SPECIAL/ SUPPLEMENTARY EXAMINATION**  
**FOR THE DEGREE OF BACHELOR SCIENCE**

**COURSE CODE:** STA 344

**COURSE TITLE:** REGRESSION ANALYSIS AND ANOVA

**DATE:** 19/09/17

**TIME:** 3 PM -5 PM

---

**INSTRUCTIONS TO CANDIDATES**

Answer Question One and Any other TWO Questions

TIME: 2 Hours

This Paper Consists of 4 Printed Pages. Please Turn Over.



**QUESTION 1:**

The ranks of 12 students according to their marks in mathematics and statistics were as follows:

Student no.	1	2	3	4	5	6	7	8	9	10	11	12
Mathematics	5	2	1	6	8	11	12	4	3	9	7	10
Statistics	4	3	2	7	6	9	10	5	1	11	8	12

- (i) Obtain the rank correlation coefficient and hence comment on the students' performance Mathematics and Statistics. (7 marks)
- (ii) Fit a simple linear regression model to the data values. What do you conclude about the expected performance in mathematics in relation to that in Statistics? (8 marks)
- (iii) Do the above results in (i) and (ii) lead to the same logical conclusion on the performance of Students in the two subjects?.Discuss. (5 marks)

(c) In an experiment, welding fluxes with differing chemical compositions were prepared. Several welds using each flux were made. The results of hardness measurements of five Welds using each of the four fluxes are presented below.

Flux	Sample values	Sample mean
A	250 264 256 260 239	253.8 ( $\mu_1$ )
B	263 254 267 265 267	263.2 ( $\mu_2$ )
C	257 279 269 273 277	271.0 ( $\mu_3$ )
D	253 258 262 264 273	262.0 ( $\mu_4$ )

Can we conclude that there are differences in the population means among the four flux types?. That is, test

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad \text{versus}$$

$$H_1: \text{not all the means are equal at 10\% level of significance.}$$

(10 marks)

**QUESTION 2:**

In the estimation of regression equations of two variables X and Y, the following results were Obtained:

$$\bar{x} = 90 \quad \bar{y} = 70 \quad N = 10$$

$$\sum x^2 = 6360 \quad \sum y^2 = 2860 \quad \sum xy = 3900$$



QUESTION 4:

A modulation study on the yield of a certain plant gave the following data

Dry weight of plants(mg), Y	Root length(cm), $X_1$	Shoot length (cm ), $X_2$
412	28.7	21.5
226	13.4	11.7
292	14.6	12.9
323	18.0	14.8
233	12.1	11.0
368	23.4	19.2
239	12.6	11.4
382	30.2	22.6
218	11.6	10.8
222	12.0	10.2
214	12.4	10.1

- (a) Fit a linear regression equation of Y on  $X_1$  and  $X_2$ . ( 12 marks )
- (b) Test the statistical significance of each partial regression coefficient ( 8 marks )

QUESTION 5:

The removal of ammoniacal nitrogen is an important aspect of treatment of Leachate at landfill sites. The rate of removal (in percent per day) is recorded for several days for each of several treatment methods. The results are presented in the following table.

Treatment	Rate of Removal			
A	5.21	4.65		
B	5.59	2.69	7.57	5.16
C	6.24	5.94	6.41	
D	6.85	9.18	4.94	
E	4.04	3.29	4.52	3.75

- (a) Construct an ANOVA table. What is the F-value in this case? ( 15 marks )
- (b) Can you conclude that the treatment methods differ in their rates of removal? ( 5 marks )