

**KABARAK**



**UNIVERSITY**

**EXAMINATIONS**

**2008/2009 ACADEMIC YEAR**

**FOR THE DEGREE OF BACHELOR OF SCIENCE IN  
ECONOMICS AND MATHEMATICS**

**COURSE CODE: MATH 328**

**COURSE TITLE: REGRESSION ANALYSIS AND ANALYSIS  
OF VARIANCE**

**STREAM: Y3S2**

**DAY: THURSDAY**

**TIME: 9.00 – 11.00 A.M.**

**DATE: 25/03/2009**

---

**INSTRUCTIONS:**

Answer questions **ONE** and any other **TWO** questions.

**PLEASE TURN OVER**

**QUESTION ONE (20 MARKS)**

- a) Define the following terms;
- i) Correlation **(2 marks)**
  - ii) Coefficient of multiple determination **(3 marks)**
  - iii) Partial correlation coefficient **(3 marks)**
- b) Write the equation of the multiple regression linear model in exact linear form and stochastic form for the case of;
- i) 4 independent or explanatory variables **(4 marks)**
  - ii) k independent or explanatory variables **(4 marks)**
  - iii) Why would you expect most observed values of Y not to fall exactly on a straight line **(2 marks)**
- c) For purpose of estimating this years inventory a computer company samples 6 dealers in each case getting figures for both this years and last years

Inventory last years(X)	Inventory this year(Y)
70	60
260	320
150	230
100	120
20	50
60	60

- i) Calculate the least squares regression equation for this year's inventory level against last years and plot the regression line on a scatter diagram **(8 marks)**
- ii) Calculate correlation coefficient **(4 marks)**

### **QUESTION TWO (20 MARKS)**

The annual sales revenue (in Ksh) of a product as a function of sales force (Number of salesperson) and annual advertising expenditure (in Ksh) for the past 10 years are summarized below;

Annual sale revenue (Y)	Sales force ( $X_1$ )	Annual Advertising expenditure ( $X_2$ )
20	8	28
23	13	23
25	8	38
27	18	16
21	23	20
29	16	28
22	10	23
24	12	30
27	14	26
35	20	32

Design a regression model to forecast the annual sales revenue of the product using matrix method.

### **QUESTION THREE (20 MARKS)**

Five doctors each test treatments for a certain disease and observe the number of days each patient takes to recover. The results are (recovery time in days) given below;

Doctors	Treatments				
	1	2	3	4	5
1	10	14	23	18	20
2	11	15	24	17	21
3	9	12	20	16	19
4	8	13	17	17	20
5	12	15	19	15	22

Discuss the difference between doctors and treatments at 5% level of significance.

**QUESTION FOUR (20 MARKS)**

The table below gives the per capita GDP to the nearest \$100(Y) and the percentage of the economy represented by Agriculture ( $X_1$ ) and the male literacy rate ( $X_2$ ) reported by World Bank Development Indicators for 1999 for 15 Latin Americans countries.

Country	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$Y_i$	76	10	44	47	23	19	13	19	8	44	4	31	24	59	37
$X_1$	6	16	9	8	14	11	12	10	18	5	26	8	8	9	5
$X_2$	97	92	85	96	91	83	93	81	74	93	67	92	94	97	93

- (i) Find the least squares regression equation of Y on ( $X_1$ ) and ( $X_2$ ) and interpret the results **(10 Marks)**
- (ii) Test at the 5% level for the statistical significance of the slope parameters **(10 marks)**

**QUESTION FIVE (20 MARKS)**

The table given below reports the aggregate consumption Y in billions and disposal income X in billions for a developing economy for 12 years

Year	n	$Y_i$	$X_i$
1988	1	102	114
1989	2	106	118
1990	3	108	126
1991	4	110	130
1992	5	122	136
1993	6	124	140
1994	7	128	148
1995	8	130	156
1996	9	142	160
1997	10	148	164
1998	11	150	170
1999	12	154	178

- i) Draw a scatter diagram for the above data **(2 marks)**
- ii) Find the regression equation **(6marks)**
- iii) Plot the regression line and show the deviations of actual values from the estimated values **(2marks).**
- iv) Test at the 5% level of significance for the statistical significance for the parameters **(6marks).**
- v) Find the correlation coefficient **(2marks)**
- vi) Find the coefficient of determination **(2marks)**