MACHAKOS UNIVERSITY

University Examinations for 2024/2025 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES
DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCES

THIRD YEAR FIRST SEMESTER EXAMINATION FOR DEGREE IN BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING BACHELOR OF EDUCATION (SCIENCE)

SST 300: ECONOMETRICS I

DATE: 23RD JANUARY, 2025 **TIME: 2.00 PM – 4.00 PM**

INSTRUCTIONS:
*Answer question **ONE (Compulsory)** and **any other TWO** questions*

QUESTION ONE *(COMPULSORY)* (30 MARKS)

A company wants to predict its monthly profit (Y) based on two factors; employee training hours (X_1) and product defect rate (X_2). The company collects the following data for 12 months:

Month	Training Hours (X_1)	Defect Rate (X_2)	Profit (Y)
1	6	5	20
2	6	6	24
3	4	2	21
4	6	3	19
5	4	2	21
6	7	3	19
7	8	2	22
8	8	5	20
9	7	7	20
10	4	4	21
11	4	4	25
12	8	5	20

a) Fit a multiple regression model relating profit to the two independent variables **(8 marks)**

b) Test for significance of regression. What conclusions can you draw? **(6 marks)**

c) Calculate R^2 and R^2_{adj} for this model. Comment on the findings. **(5 marks)**

d) Use t tests to assess the contribution of each regressor to the model. Explain your findings. **(6 marks)**

e) Find a 95% confidence interval for β_1 and β_2 **(3 marks)**

f) Find a 95% prediction interval for a new observation on profit when $x_1 = 5.0$ and $x_2 = 8.0$ **(2 marks)**

QUESTION TWO (20 MARKS)

An economist studying the relation between household weekly electricity consumption (Y) and number of rooms in the home (X) obtained the following data:

i	1	2	3	4	5	6	7	8
x_i	2	4	5	3	7	6	6	3
y_i	63	67	74	64	75	69	90	60

a) Assuming that the normal error simple linear regression model is appropriate, obtain the estimated regression. **(4 marks)**

b) Plot the residuals e_i against x_i. What does the residual plot show? **(4 marks)**

c) Omit case 7 from the data and obtain the estimated regression function based on the remaining seven cases. **(3 marks)**

d) Compare the new estimated regression function to that obtained in part a). What can you conclude about the effect of case 7? **(3 marks)**

e) Using the new fitted regression function, obtain a 99 percent prediction interval for a new Y observation at $X = 12$. **(3 marks)**

f) Does observation y_7 fall outside this prediction interval? What is the significance of this? **(3 marks)**

QUESTION THREE (20 MARKS)

a) Explain the following violations of the classical linear regression model, describing how they can be identified, their effects and remedial actions that can be taken.

i. Multicollinearity **(5 marks)**

ii. Autocorrelation **(5 marks)**

b) A staff analyst for a manufacturer of microcomputer components has compiled monthly data for the past 16 months on the value of industry production of processing units that use these components (X, in million KESs) and the value of the firm's components used (Y. in thousand KESs). The analyst believes that a simple linear regression relation is appropriate but anticipates positive autocorrelation. The data follows:

X_t	10.3	10.2	10.1	9.8	9.7	9.4	9.8	10.2	10.5	10.7
Y_t	2.1	2.0	1.8	2.0	1.9	1.9	2.0	2.1	2.2	1.9

i. Fit a simple linear regression model **(4 marks)**

ii. Conduct the Durbin - Watson test for autocorrelation using $\alpha = 0.05$ **(6 marks)**

QUESTION FOUR (20 MARKS)

a) Describe the Brown-Forsythe and Breusch-Pagan tests for the constancy of the error variance **(6 marks)**

b) The data below was used to study the relationship between income and years of education:

Income (KES 10,000)	2.5	3.0	3.2	3.6	2.8	5.7	4.7	5.5	7.5	9.5
Education (Years)	8	10	12	11	15	13	14	16	18	20

i. Fit a linear regression model **(4 marks)**

ii. Conduct the Brown-Forsythe test by dividing the data into two equal groups **(5 marks)**

iii. Perform the Breusch-Pagan test on the data and make relevant conclusion **(5 marks)**

QUESTION FIVE (20 MARKS)

a) The data below contains the trade openness and county GDP growth of five counties over a period of 3 years. An analyst would like to model county GDP growth as a function of trade openness (fixed effects) while accounting for differences between counties as random effects.

County	Trade Openness (%)	GDP Growth (%)
A	58	4.5
A	37	1.6
A	68	4.3
B	36	1.8
B	62	4.4
B	37	2.1
C	72	4.1
C	63	3.2
C	58	5.3
D	42	1.9
D	36	2.2
D	42	2.3
E	38	3.3
E	57	4.8
E	44	2.2

Estimate the fixed and random effects

(20 marks)