MAASAI MARA UNIVERSITY

REGULAR UNIVERSITY EXAMINATIONS
2023/2024 ACADEMIC YEAR
FIRST YEAR FIRST SEMESTER

SCHOOL OF BUSINESS AND ECONOMICS
MASTER SCIENCE IN ECONOMICS AND STATISTICS

COURSE CODE: ECS 8122
COURSE TITLE: STATISTICAL METHODS AND THEORY

DATE:2/2/2024 **TIME: 1430-1730 HRS**

INSTRUCTIONS TO CANDIDATES

*Answer question **ONE (compulsory)** and any other **TWO** questions.*

Question One (30 Marks)

a) Suppose the relationship between two variables X_j and Y_j can be given by the model $Y_j = \alpha + \beta X_j + e_j$, $j = 1, 2, ...n$, where $(Y_j, X_j) \in \Re x \Re$, Y_j is a dependant variable, X_j is random and independent of e_j , and $e_j's$ are random error variables. Using appropriate assumptions, derive the estimators $\hat{\alpha}$ and $\hat{\beta}$ using Least Squares Method (8 marks)

b) Suppose the relationship between two variables X_j and Y_j can be given by the model $Y_j = \alpha + \beta X_j + e_j$, $j = 1, 2, ...n$, where $(Y_j, X_j) \in \Re x \Re$, Y_j is a dependant variable, X_j is random and independent of e_j , and $e_j's$ are random error variables. If a researcher uses the model and least squares method for prediction;

(i) State the assumptions that he would make (3 marks)

(ii) Suppose the researcher observes X_{n+1} so that the forecast of Y_{n+1} is $\hat{Y}_{n+1} = \hat{\alpha} + \hat{\beta} X_{n+1}$,

where the OLS estimators $\hat{\alpha}$ and $\hat{\beta}$ are calculated on the basis of the observations for $j = 1, 2, ..., n$. Find an expression for the variance of the forecast error. (4 marks)

c) The following are the average income per acre produced by a commercial game rancher over a ten year period

Time(X)	1	2	3	4	5	6	7	8	9	10
Income per acre{in KShs '000(Y)}	510	590	640	740	780	890	1010	1070	1100	1180

(i) Fit the Least Squares Estimates of the Parameters (3 marks)

(ii) Find the Fitted or Predicted Line (2 marks)

d) A production plant has two fabricating systems: one uses automated equipment, the other is manually operated. Since the automated system costs more to install, we wish to know whether it provides increased production in terms of the mean number of finished products fabricated per day. Consider the accompanying data below on the production for 32 days.

Automated	56	36	67	88	86	97	18	57	84
	28	48	28	25	42	47	26	51	
Manual	42	22	28	22	37	52	21	44	18
	22	42	37	16	51	41			

Test the hypothesis that the daily mean of production for the automated system is higher than the manual one. (5 marks)

Question Two (20 Marks)

a) The weight of a certain product has a mean weight of 38kg. An observation was made that in products of a given firm were underweight. As evidence, officials sampled 20 of such products from the firm. Test an appropriate hypothesis at $\alpha = 0.05$.

31.9	42.3	25.2	37.4	40.4	31.1	29.2	31.7	31.6	46.4
42.8	40.5	32.7	29.4	34.5	41.5	46.6	33.9	42.4	34.5

(10 marks)

b) It is hypothesized that the location of a business affects its sales. A small company has branches in two locations trading on similar items. The given data are the daily sales in thousands of KES. Test the hypothesis of different mean sales at the two sites. (10 marks)

Day	1	2	3	4	5	6	7	8
Location A	5.51	7.14	5.45	3.95	7.74	5.22	4.39	4.55
Location B	4.71	5.47	6.12	3.69	5.43	5.20	2.16	4.34

Day	9	10	11	12	13	14	15	16
Location A	5.74	6.20	3.29	6.62	4.01	5.55	4.60	7.08
Location B	3.17	4.44	4.99	2.11	4.17	5.15	6.13	4.09

Question Three (20 Marks)

a) The data below represents the education level and income of employees in a large organization. Test the independence of the two variables

Education level	Income (KES)		
	< 20000	20000 – 50000	> 50000
KCSE	54	44	10
Diploma	25	25	32
Bachelors	45	33	36
Post graduate	12	36	48

(6 marks)

b) The following are the daily stock prices of a company listed at the Nairobi Stock Exchange during the month of September, 2017

17.2	17.6	15.4	18.7	16.9	19.5	19.1	18.9	20.3	19.5	18.6
16.8	19.6	22.4	17.8	20.0	18.5	22.4	17.5	18.5	19.6	

i. Fit a three period moving average model to the data (4 marks)

ii. Fit a linear trend model to the data (6 marks)

iii. Compare the adequacy of fit for the two models in i. and ii. above (4 marks)

Question Four (20 Marks)

a) Four different strategies of marketing were analyzed to determine their effectiveness in boosting sales at various locations. The data are shown below. Perform an appropriate test at $\alpha = 0.05$.

Strategy	Percentage Increase of Sales				
A	13.5	13.4	14.1	14.2	
B	13.2	12.7	12.6	13.9	
C	16.8	17.2	16.4	17.3	18.0
D	18.1	17.2	18.7	18.4	

(8 marks)

b) The birth weights kilograms of 30 children were recorded as follows:

2.0	2.1	2.3	3.0	3.1	2.7	2.8	3.5	3.5	3.7
4.0	2.3	3.5	4.2	3.7	3.2	2.7	2.5	2.7	3.8
3.5	3.0	2.6	2.8	2.9	3.5	4.1	3.9	2.8	2.7

i. Present the data in a grouped frequency distribution table (6 marks)

c) Mean morphometic values of workers of Apis Cerana measured in mm were as follows at thirteen locations:
Labrum length: 0.42, 0.43, 0.40, 0.33, 0.36, 0.40, 0.32, 0.40, 0.38, 0.36, 0.40, 0.33, 0.35
Labrum breadth: 1.10, 1.02, 1.00, 1.00, 1.02, 0.99, 1.00, 1.00, 1.00, 1.00, 1.01, 1.00, 0.99

i. Draw the scatter diagram (2 marks)

ii. Find the regression of labrum breadth on labrum length (2 marks)

iii. Estimate the labrum breadth for labrum length =0.5 (2 marks)

iv. Calculate the correlation between labrum breadth and length (2 marks)

Question Five (20 Marks)

The production levels of a finished product (produced from sheets of stainless steel) have varied quite a bit, and management is trying to devise a method for predicting the daily amount of finished product. The ability to predict production is useful for scheduling labor, warehouse space, and shipment of raw materials and also to suggest a pricing strategy. The number of units of the product (Y) that can be produced in a day depends on the density (X_1) of the sheets being processed, and the tensile strength of the steel (X_2). The data are taken from 20 days of production.
The observations are given in the table below. Fit a multiple linear regression model to the data and comment on the goodness of fit.

(20 marks)

Day	Unit produced	Density	Tensile strength
1	47	3.5	1.6
2	33	4.0	1.9
3	29	5.8	1.9
4	29	7.1	2.3
5	36	3.3	1.6
6	33	7.9	2.3
7	52	2.0	1.3
8	50	1.9	1.9
9	31	5.6	1.1
10	39	2.6	1.4
11	44	4.4	1.5
12	34	9.0	1.2
13	42	4.3	1.5
14	46	4.2	1.4
15	32	5.5	1.4
16	34	6.0	1.7
17	27	6.2	2.1
18	46	2.6	2.1
19	33	6.0	1.1
20	45	2.7	1.8

/END/