MACHAKOS UNIVERSITY

University Examinations for 2022/2023 Academic Year

SCHOOL OF BUSINESS, ECONOMICS AND HOSPITALITY AND TOURISM MANAGEMENT

DEPARTMENT OF ECONOMICS

THIRD YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF ECONOMICS AND FINANCE

BACHELOR OF ECONOMICS AND STATISTICS

BACHELOR OF ECONOMICS

EES 301: STATISTICS FOR ECONOMISTS II

DATE: **TIME:**

INSTRUCTIONS

- **(i) Answer question one (Compulsory) and any other two questions**
- **(ii) Do not write on the question paper**
- **(iii) Show explicitly all formulas and calculations**
- **(iv) Non – programmable calculators may be used.**
- **(v) Note that there is no borrowing in the exam room; a candidate MUST have his/her own stationeries.**

QUESTION ONE (COMPULSORY)

a) Suppose that a researcher using data on Family income (FI) and average family consumption from 15 Kenyan families estimates the following OLS regression using STATA 13

$C = \beta_0 + \beta_1 FI$ and obtained the following output.

```
. regress consumption income
```

Source	SS	df	MS	Number of obs = 15
Model	3351406.23	1	3351406.23	F(1, 13) = 8144.59
Residual	5349.35306	13	411.488697	Prob > F = 0.0000
				R-squared = 0.9984
				Adj R-squared = 0.9983
Total	3356755.58	14	239768.256	Root MSE = 20.285

consumption	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	.706408	.0078275	90.25	0.000	.6894978	.7233182
_cons	-184.0779	46.26183	-3.98	0.002	-284.0205	-84.13525

i) Write the estimated regression equation for the full model, the consumption function (2 marks)

ii) Carry out a test for the null hypothesis. Ho: $\beta_1 = 0$. State the test statistic, the P-Value, and your conclusion about whether or not to reject the null hypothesis using α =0.05 (2 marks)

iii) Explain in words in the context of Keynesian consumption function what the hypothesis test in (ii) above mean (2 marks)

iv) If you were to fit the simple regression model using family income as the only predictor of family consumption. Would the results be that family income is a significant predictor of family consumption? explain how you would know using the information provided in the output (2 marks)

v) Construct a 95% confidence interval for the slope coefficient (β_1) (2 marks)

vi) Based on the results for (i) to (v), comment and interpret the results of the estimated model. Are the results consistent with Keynesian consumption theory? (3 marks)

vii) Explain the importance of hypothesis testing in econometric analysis and highlight the circumstances under which a researcher would use a *t*-test other than the *Z* test (3 marks)

b) The makers of Duracell batteries want to demonstrate that their size AA battery lasts an average of at least 45 minutes longer than Duracell's main competitor, the Energizer. Two independent random samples of 100 batteries of each kind are selected. The sample average lives for Duracell and Energizer batteries are found to be $\bar{X}_1 = 308$ minutes and $\bar{X}_2 = 254$ minutes respectively. Assume $\sigma_1 = 84$ minutes and $\sigma_2 = 67$ minutes. Is there any evidence to substantiate Duracell's claim that its batteries last, on average, at least 45 minutes longer than Energizer of the same size? (5 marks)

c) The following 3x3 contingency table contains observed frequencies for a sample of 240. Test the independence of the row and column variables with $\alpha = 0.05$ (6 marks)

	Column variable		
Row variable	A	B	C
P	20	30	20
Q	30	60	25
R	10	15	30

d) Distinguish between Type I error and Type II error (3 marks)

QUESTION TWO (20 MARKS)

a) A newly married couple in December 2022 is planning to give birth and raise three children in the next five years. The couple is interested as matter of chances on the number of boys out of the 3 births.

i) Generate a sample space of all possible outcomes from the three births (3 marks)

ii) Generate a probability distribution of the number of boys out of the three births (3 marks)

iii) Determine the mean and variance of the probability distribution in (ii) above (3 marks)

b) A firm conducts a study to determine if the absenteeism of day workers is different from those of employees who work the night shift. A comparison is made of 150 workers for each shift. The results show that 37 day workers have been absent at least five times over the past year, while 52 night workers have missed at least five times. What does this reveal about the tendency for absenteeism among the workers? Calculate a 90% C.I for the difference in the proportion of workers on the two shifts who missed at least five times. (5 marks)

c) At a parking place, the average number of car arrivals during a specified period of 15 minutes is 2. If the arrival process is well described by a Poisson process, find the probability that during a given period of 15 minutes

i) At least two cars will arrive (2 marks)

ii) At most three cars will arrive (2 marks)

iii) Between 1 and 3 cars will arrive (2 marks)

QUESTION THREE (20 MARKS)

a) As more Kenyans seek to escape from urban pressures, the burden on our parks has shown a marked increase in the number of weekend campers. Lion magazine recently reported that Machakos Peoples park hired a third-year Economist from MKSU to study the financial position of the park. Part of his effort required the comparison of park revenues from various sources, including camping fees, hiking licenses and canoe rentals. Displayed here are the data for several randomly selected visitors. Determine if there is a difference in the park's mean revenues from these three activities.

Visitor	Camping	Hiking	Canoeing
1	$47	$30	$19
2	32	18	25
3	35	27	20
4	25	35	22
5	38	-	25
6	35	-	-

(12 marks)

b) Use relevant examples to Distinguish between Poisson Hypergeometric and exponential distributions and explain the situations where each distribution may be used. (6 marks)

c) Explain the meaning of P-Value in tests of hypothesis (2 marks)

QUESTION FOUR (20 MARKS)

a) Use a relevant example to distinguish between an estimate and an estimator (4 marks)

b) A market research firm in Nairobi collected data on the daily prices and quantities of a product traded for 12 days, as shown by the following data.

Day	***Quantity (Y)***	***Price(X)***
1	*69*	*9*
2	*76*	*12*
3	*52*	*6*
4	*56*	*10*
5	*57*	*9*
6	*77*	*10*
7	*58*	*7*
8	*55*	*8*
9	*67*	*12*
10	*53*	*6*
11	*72*	*11*
12	*64*	*8*

Required

i) Develop the estimated regression equation that could be used to describe the relationship between the Prices and the quantities of the product in the market (8 marks)

ii) Is it a demand or a supply function? Explain your answer (3 marks)

Do you believe the estimated regression equation would better predict the quantities traded in the market? use r^2 to support your answer (4 marks)

c) Distinguish between the One-tailed test and the Two-tailed test (2 marks)

QUESTION FIVE (20 MARKS)

a) Fertile areas of the Nyanza region had relatively high population growth rates during the 1990s. Data were collected from residents living in the fertile and as well non-fertile locations of Nyanza. Assume that the following sample results were obtained on the ages of individuals in two populations.

Fertile Areas **Non-Fertile areas**

$n_1 = 150$ $n_2 = 175$

$\bar{X}_1 = 39.3\ years$ $\bar{X}_2 = 35.4\ years$

$S_1 = 16.8\ years$ $S_2 = 15.2\ years$

Test the hypothesis of no difference between the two population means. Use α= 0.05

(6 marks)

b) In an opinion survey regarding a certain political issue, there was some question as to whether or not the eligible voters under 25 years of age might view the issue differently from those over 25 years. 1500 individuals of those over 25 years were interviewed, and 1000 of those under 25 years were interviewed with the following results

	Opposed	Undecided	Favour	Total
Under 25	400	100	500	1000
Over 25	600	400	500	1500
Total	1000	500	1000	2500

Test the null hypothesis that there is no evidence of a difference of opinion due to the different age grouping; take α = 0.05 (10 marks)

c) State four properties of the standard normal distribution (4 marks)