MACHAKOS UNIVERSITY

University Examinations for 2020/2021

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

THIRD YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF ENVIRONMENT STUDIES (ENVIRONMENTAL RESOURCE AND CONSERVATION)

BACHELOR OF ENVIRONMENT SCIENCE

ESU 302: STATISTICS

DATE: 10/8/2021 **TIME: 8:30 -10:30 AM**

INSTRUCTION: *Answer Question **ONE** which is compulsory and any other **TWO** Questions*

QUESTION ONE (COMPULSORY) (30 MARKS)

a) Differentiate the following terms as they apply in scientific research

i) Sample and a population

ii) Response variables and independent variables

iii) Descriptive statistics and inferential statistics

iv) Null hypothesis and Alternative hypothesis

v) Significance level and confidence level in hypothesis stating (10 marks)

b) Two science based methods for restoring soil productivity were separately used and the increased yield in bags per acre recorded as summarized in the table below.

Methods	Increased yield (bags) per acre						
Method I	8	12	13	9	3	-	-
Method II	10	8	12	15	6	8	11

(i) Determine whether the two methods effect differed significantly , at $\alpha = 5\%$

(ii) Test the hypothesis that the increased yield for method II was 11.9 $\alpha = 5\%$

(10 marks)

c) (i) Determine the regression equation of Y on X given that

$$\sum xy = 130,\ \sum x^2 = 2400,\ \bar{x} = 60 \text{ and } \bar{y} = 4.$$

(ii) Predict the value of Y given $X = 50$ (6 marks)

(iii) Highlight four assumptions for regression modeling (4 marks)

QUESTION TWO (20 MARKS)

a) The data below is a summary of the slim possible participants weights in kilograms before and after the season.

Participant	A	B	C	D	E	F	G	H
Weight Before	85	124	172	123	111	139	99	77
Weight after	105	160	175	143	156	127	95	100

Test the hypothesis that on average the exercise did not result to any significant weight loss at 10% significance level (8 marks)

b) Below was the daily harvest of a tropical plant during a particular week in the year 2017.

Harvested fruits	11-20	21-30	31-40	41-50	51-60
Frequency	3	6	11	3	2

Determine;

i) Average harvest per day
ii) Median
iii) Mode and
iv) Standard deviation. (12 marks)

QUESTION THREE (20 MARKS)

a) The information relates to health records from County XYZ randomly selected

Gender	Males	Females
Average lifespan	50	58
Standard deviation	12	9
Sample size	100	150

Determine;

i) The gender with the higher dispersion lifespan

ii) The combined standard deviation (10 marks)

b) Using the data below determine if there was any correlation between the sodium concentration and Opuntia area coverage. (10 marks)

Sodium conc.	1.5	5.0	5.0	4.5	7.0	5.5	9.0	6.5	7.5	10.5	11.0	11.5
Opuntia area	14.0	30.0	71.0	77.5	75.5	121.5	132.0	159.0	181.5	184.0	194.5	219.0

QUESTION FOUR (20 MARKS)

a) Discuss any two test statistics and state when appropriate to apply them. (4 marks)

b) Discuss the four levels of measurement scales (8 marks)

c) The table below shows the scores of eight interviewees for an environmental conservationist job on aptitude and subject matter knowledge.

Interviewee	A	B	C	D	E	F	G	H
Aptitude test score	15	20	28	12	40	60	20	80
Subject matter score	40	30	50	30	20	10	30	60

Compute the rank coefficient of correlation (8 marks)

QUESTION FIVE (20 MARKS)

a) The table below shows the age of savanna antelope and the weight in kilograms ;

Age(Years)	9	7	5	10	4	5	3	2
Weight (Kg)	45	42	41	60	30	34	25	20

i) Use least square methods to calculate regression equation of weight on age

ii) Predict Y when X is 8 and 20 (10 marks)

b) Highlight the difference in the following terms as applied in statistics

i) Stratified sampling and snowballing sampling methods

ii) Sampling errors and Non-sampling errors

iii) Data skewness structure and data kurtosis structure

iv) Type I error and Type II error in test of hypothesis

v) Continuous variables and discrete variables (10 marks)