MACHAKOS UNIVERSITY

University Examinations for 2024/2025 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES
DEPARTMENT OF MATHEMATICS AND STATISTICS

THIRD YEAR SECOND SEMESTER EXAMINATION FOR DEGREE IN BACHELOR OF SCIENCE IN ENVIRONMENTAL SCIENCES

ESU 302/307: APPLIED STATISTICS

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

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

QUESTION ONE (30 MARKS)

(a) Differentiate the following terms as they apply in statistics

i. Discrete random variable and continuous random variable

ii. Descriptive statistics and inferential statistics

iii. Null hypothesis and alternative hypothesis (6 marks)

(b) A researcher wished to confirm the conformity to product composition where the element of interest was supposed to be 6mg. A randomly selected sachets had the following weights of the element : 6, 5, 9, 7, 6, 1, 0, 3, 4, and 8 . The t-test output was as given

The sample t-test output, for the test value=65

	t	df	Sig.(2-tailed)	Mean difference	95% CI of the difference	
					Lower	Upper
Weight	-1.19	9	0.2645	4.500	2.8090	6.9911

i. Construct the hypothesis statements for the above scenario (3 marks)

ii. Based on the t-value and the sig.(2-tailed) value make statistical conclusion (4 marks)

(c) During the last week to the University examination the office of registrar, academic receives an average of 20 students per hour

Determine the probability that the office receives

i. No student in the next one hour

ii. More than 3 students in the next one hour (4 marks)

(d) The probability that John hits the bull eye in a dart game is 0.12. Determine the probability that in ten trials he hits it;

i. At least once

ii. Exactly 4 times (4 marks)

(e) The probability mass function of a discrete random variable X is given below.

X	40	80	120	150	200
$P(X=x)$	*a*	0.20	0.25	*b*	0.15

Suppose $P(X<150)=0.69$. Determine

i. The value of *a* and *b* (3 marks)

ii. E(X) (3 marks)

iii. Var(X) (3 marks)

QUESTION TWO (20 MARKS)

a) Differentiate the following terms as they apply in test of hypothesis

i. Type I error and type II error

ii. One Tailed and Two Tailed Hypothesis (4 marks)

b) The mean and standard deviation for the heights of a randomly selected sample of 36 plants within the university compound are calculated to be 2.6m and 0.3m, respectively. Determine 95% confidence interval for the mean of the entire tree coverage. (4 marks)

c) A sample of 200 persons were randomly selected in a wellness clinic with cardiovascular disease, 100 had hypertension as summarised below:

		Hypertension	
		Yes	No
Cardiovascular	Yes	65	55
	No	35	45

Using the data above;

i. Calculate Odds Ratio (4 marks)

ii. Tests whether there is any significant relationship between hypertension and cardiovascular heart attack at $\alpha = 0.01$. (8 marks)

QUESTION THREE (20 MARKS)

a) Highlight four characteristics of a binomial distribution (4 marks)

b) A large group of students pursuing BSc in environmental final test was normally distributed with an arithmetic mean of 60 marks and a standard deviation of 15 marks.

i. If a student is randomly chosen from the class, determine the probability that she/he scored between 45 marks and 75 marks (4 marks)

ii. If the pass mark was 40, what percent of the students passed the examination (2 marks)

iii. If the top 10% of the students were to participate in a tree growing conference, in Rwanda. What is the minimum marks would a student score to participate? (4 marks)

c) A health facility with Poisson patients-arrival stream had an average of 5 patients arrive per hour. Determine the probability that in any one-hour interval

i. No patient arrives

ii. At most two patients arrive

iii. More than two patients arrive (6 marks)

QUESTION FOUR (20 MARKS)

a) Distinguish between one- way ANOVA and Kruskal Wallis test statistics as they apply in data analysis (4 marks)

b) According to the last Kenya National Bureau of Statistics Census records, one of the counties recorded that; 63.9% of the population is married, 7.7% widowed, 6.9% divorced (and not remarried), and 21.5% single (never been married). Seven years after the census a sample of 500 adults from the same county showed that 310 were married, 40 widowed, 30 divorced (not remarried), and 120 single (Never married). At the 0.05 significance level can it be concluded that the sample findings are different from the census results? (6 marks)

c) ANOVA table output below shows the data results after analysing the effect of different fertilizers on the yield tried in different plots.

Fertilizer	**Yield Per acre**				
1	26	31	24	30	24
2	16	15	13	21	15
3	18	22	20	16	24

The ANOVA analysis was conducted and the output is partially completed in the table below;

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F-critical*
Between Groups	310	2	*x*	*y*	0.00047236	*z*
Within Groups	120	12	10			
Total	430	*w*				

i. Construct the hypothesis statement for the analysis. (2 marks)

ii. Complete the values w, x, y and z in the ANOVA table (4 marks)

iii. Interpret the finding in this analysis at $\alpha = 0.01$ (4 marks)

QUESTION FIVE (20 MARKS)

a) Differentiate between each of the following terms as they apply in scientific research:

i. Sample statistic and Population parameter

ii. Point estimation and interval estimation

iii. Parametric test statistics and Non-parametric test statistics

iv. Primary data source and secondary data source

v. One way ANOVA and two way ANOVA (12 marks)

b) After the National Tree Growing Restoration Campaign, a random sample to check the trees survival rate in two different ecological zones (A and B) was conducted and given below is part of the analysis output.

	A	*B*
Mean	19	20
Variance	15.5	9
Observations	5	7
Pooled Variance	11.6	
Hypothesized Mean Difference	0	
df	10	
t Stat	-0.501434723	
P(T<=t) one-tail	0.313459622	
t Critical one-tail	1.812461123	
P(T<=t) two-tail	0.626919244	
t Critical two-tail	2.228138852	

Using the information in the table;

i. Construct the hypothesis statements for the above scenario (3 marks)

ii. Based on the t-value make statistical conclusion (5 marks)