MANICALAND STATE UNIVERSITY OF APPLIED SCIENCES

FACULTY OF ENGINEERING, APPLIED SCIENCES & TECHNOLOGY

DEPARTMENT OF APPLIED STATISTICS

MODULE: INTRODUCTION TO STATISTICS

CODE: ASTA 101

SESSIONAL EXAMINATIONS

DURATION: 3 HOURS

EXAMINER: MR M. HAZVINANDAWA

INSTRUCTIONS

1. *Answer **ALL** questions from Section A*
2. *Answer any **three** questions from Section B*
3. *Total Marks: 100*

REQUIREMENTS

Statistical tables; Graph paper
Non-programmable scientific calculator

SECTION A: 40 MARKS

Answer **All** Questions in this Section

A1 Define the following terms as used in Statistics:

a) quantitative data;

b) parameter;

c) sample;

d) qualitative data.

[2, 2, 2, 2]

A2 A student finds that the average number of amoebas in 10 ml of pond water from a particular pond is four. Assume that the number of amoebas follows a Poisson distribution.Find the probability that in a 10 ml sample:

a) there are exactly 5 amoebas;

b) there are no amoebas;

c) there are fewer than three amoebas.

[3, 3, 4]

A3 a) In the mass production of bolts it is find that 5 percent of bolts are defective. Bolts are selected at random and put into packets of ten. A packet is selected at random. Find the probability that it contains:

i) three defective bolts;

ii) At least 8 defective bolts;

iii) less than three defective bolts.

[2, 3, 3]

A4 A roulette wheel is divided into six sectors of unequal area, marked the numbers 1, 2, 3, 4, 5, 6. The wheel is spun and X is the random variable, the number on which the wheel stops. The probability distribution of x is shown in Table 1.

Table 1

X	1	2	3	4	5	6
P(X=x)	1/16	3/16	1/4	1/4	3/16	1/16

a) E(X)

c) Var(X)

d) E(X-5)

[4,6,4]

SECTION B: [60 MARKS]

ANSWER ANY THREE QUESTIONS IN THIS SECTION

B5 a) State any five assumptions of simple linear regression analysis.

b) State and explain how to carry out residual tests and assess linear regression assumptions using each of the following;

i) scatter plots

ii) normal probability plot and quantile plots (Q-Q plots).

c) Give advantages of using a sample as opposed to using a population.

d) State and explain how missing data is handled in Statistics.

[5, 3, 3, 3, 6]

B6 a) The members of a sports team are interested in whether the weather has an effect on their results. They play 50 matches, with the following results shown on Table 2.

Table 2

	Good	Bad	Total
Win	12	4	16
Draw	5	8	13
Lose	7	14	21
Total	24	26	50

Formulate suitable null and alternative hypotheses, and use a Chis-square test to test the claim, at the 1 percent significance level, that the weather has no effect on the team's results. State your conclusion clearly.

b) A driving school examined the results of 100 candidates who took their test for the first time. It was found that out of the 40

men, 28 passed and out of the 60 women, 34 passed. The results are shown in Table 3.

Table 3

	Pass	Fail	Total
Male	28	12	40
Female	34	26	60
Total	62	38	100

Do these results indicate, at the 5 percent significance level, a relationship between the sex of a candidate and the ability to pass the driving test at the first attempt?

[10, 10]

B7 a) Show that the least squares estimates ($\hat{\{}\beta_0\,\hat{\beta}_1$) are unbiased estimaters of
$\beta_0\, and \beta_1$ respectively.

b) Derive the variance of the estimates and show that var $\hat{\beta}_1 = \hat{\sigma^2}_0 / Sxx$

i) Var $(\hat{\beta}_0) = \hat{\sigma^2}_0\,(1/n + \bar{X^2}/Sxx)$;
where $\hat{\sigma^2}_0$ is the Mean Square Error

Sxx is the variance of the independent variable X.

N is the number of data points.

[4, 4, 6, 6]

B8 A company gathered the following information on its advertising expenditure and sales generated in thousands as follows in Table 4.

Table 4

Advertising Expenditure	Sales(thousand
20	44
24	50
36	52
30	60
16	46
34	54
40	64
30	54
22	50
38	58

a) Draw a scatter plot for the data and comment.

b) Estimate the least squares regression line.

c) Calculate Pearson's Product moment correlation coefficient and interpret your answer.

d) Calculate the coefficient of determination and interpret it.

e) Predict the sales when 25000 is budgeted for advertising.

[6, 6, 3, 3, 2]

END OF QUESTION PAPER