GWANDA STATE UNIVERSITY

FACULTY OF COMPUTATIONAL SCIENCES

DEPARTMENT MATHEMATICS AND STATISTICS

REGRESSION ANALYSIS AND ANOVA

CMS 2104

Examination Paper

SEPTEMBER 2024

This examination paper consists of 6 printed pages

Time Allowed: **3 hours**

Total Marks: **100**

Examiner's Name: **Mr. R. G. Moyo**

INSTRUCTIONS

Candidates should answer **ALL** questions in Section A and **ANY THREE** in Section B

ADDITIONAL REQUIREMENTS

Scientific calculator
Statistical tables

SECTION A (40 marks)

Answer ALL questions from this section.

A1. Define the following terms

(i) Hypothesis testing [2]

(ii) Alternative Hypothesis [2]

(iii) Probability [2]

(iv) 2-Factor factorial experiment [2]

A2. The derivation of the Least Squares Method approach involves the minimization of the function:

$$\phi(a,b) = \sum_{i=1}^{n}(a + bx_i - y_i)^2$$

(i) How do we obtain a and b to minimize $\phi(a,b)$? [2]

(ii) Determine the unknown a and b [6]

A3. On least squares regression, state the function $f(x)$ that must be selected for

(i) a linear function [2]

(ii) a quadratic function. [2]

(iii) a polynomial function. [2]

A4. Second year college students were randomly assigned to three groups to experiment with three different methods of teaching. At the end of the semester, the test was given to all the 15 students. The table below gives the scores of students in the three groups.

Method I	**Method II**	**Method III**
48	55	84
73	85	68
51	70	95
65	69	74
87	90	67

Are there any significant differences between the three teaching methods. [12]

A5. Copy and complete the general ANOVA Table for (**a-f**) [6]

Source of Variance	Sum of square	D.O.F	Mean square	F-value
Treatments -A	SSA	$k-1$	**f**	$\frac{MSA}{MSE}$
Blocks	SSB	$j-1$	$\frac{SSB}{j-1}$	**a**
b	**c**	**d**	**e**	
Total	SST	N-1		

SECTION B (60 marks)

Answer ANY THREE questions from this section.

B6. An industrial engineer is investigating the effect of four assembly methods (A, B, C, D) on the assembly time for a color television component. Four operators are selected for the study. Furthermore, the engineer knows that each assembly method produces such fatigue that the time required for the last assembly may be greater than the time required for the first, regardless of the method. That is, a trend develops in the required assembly time. To account for this source of variability, the engineer uses the Latin square design shown below:

Order of Assembly	**Operator 1**	**Operator 2**	**Operator 3**	**Operator 4**
1	C=10	D=14	A=7	B=8
2	B=7	C=18	D=11	A=8
3	A=5	B=10	C=11	D=9
4	D=10	A=10	B=12	C=14

Carry out an analysis of variance(ANOVA) to determine if there are any significant differences between the four assembly methods. Was using different operators and different assembly orders necessary?. Use $\alpha = 0.05$ [20]

B7. Use **APPENDIX A** to answer the whole of this question. **APPENDIX A** is an SPSS output of a multiple linear regression of 8 variables. The statistical analyst was determing the relationship that exist between the profit, price and quantity of selected items sold by a particular shop.

(i) Identify all dependent variables. [3]

(ii) Write down the equation of the regression line. [3]

(iii) What will happen to the overal profit(Y) when Pbatter increases by 1 unit? [2]

(iv) State the correlation coefficient and comment on the relationship that exist between the profit and its predictors. [3]

(v) Determine the significance of the regression constant in determining the profit. [2]

(vi) Determine all the independent variables that are significant in determining the profit? [5]

(vii) Identify the two items or products that does not contribute much to the profit [2]

B8. (i) Define trend analysis. [2]

(ii) Fit a multiple linear regression using the data in the table below. Use y as the dependent variable depending on x_1 and x_2. [18]

y	1.45	1.93	0.81	0.61	1.55	0.95	0.45	1.14	0.74	0.98	1.41
x_1	0.58	0.86	0.29	0.20	0.5	0.28	0.08	0.41	0.22	0.35	0.59
x_2	0.71	0.13	0.79	0.20	0.56	0.92	0.01	0.60	0.70	0.73	0.13

B9. Corrosion Fatigue in metals has been defined as the simultaneous action of cyclic stress and chemical attack on a metal structure. A widely used technique for minimizing corrosion fatigue damage in aluminum involves the application of a protective coating. A study conducted by the Departments of Mining and Metallurgical Engineering at Gwanda State University used three different levels of humidity.
Low: 20% to 25% relative humidity
Medium: 55% to 60% relative humidity
High: 86% to 91% relative humidity,
and three types of coatings
Uncoated: no coating
Anodized: sulphiric acid anodic oxide coating
Conversion: chromate chemical conversion coating.
The corrosion fatigue data, expressed in thousands of cycles to failure, were recorded as follows:

COATING	HUMIDITY LEVEL **(Low)**	HUMIDITY LEVEL **(Medium)**	HUMIDITY LEVEL **(High)**
Uncoated	361 469 466 937 1069 1357	314 522 244 739 261 134	1344 1216 1027 1097 1011 1011
Anodized	114 1032 1236 92 533 211	322 471 306 130 68 398	78 466 387 107 130 327
Conversion	130 1482 841 529 1595 754	252 874 105 755 847 573	586 524 402 751 846 529

Perform an Analysis of Variance (ANOVA) using $\alpha = 0.05$ to test for the significance of main and interaction effects. Also test for the significant differences between the 9 treatments. [20]

END OF QUESTION PAPER

APPENDIX A

Regression

[DataSet1]

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Pbatt, QTV, Qtapes, Qbread, Qbutter, Pbread, Ptapes[b]	.	Enter

a. Dependent Variable: Y

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.994[a]	.988	.988	1.819680

a. Predictors: (Constant), Pbatt, QTV, Qtapes, Qbread, Qbutter, Pbread, Ptapes

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-598.013	19.568		-30.561	.000
	Qbutter	-.001	.002	-.012	-.764	.446
	Qtapes	.006	.001	.135	6.175	.000
	QTV	.002	.001	.018	1.789	.076
	Qbread	.035	.012	.047	2.950	.004
	Ptapes	7.544	.298	.888	25.355	.000
	Pbread	.167	.024	.154	7.100	.000
	Pbatter	-.755	.110	-.192	-6.888	.000

a. Dependent Variable: Y