NAMIBIA UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Faculty of Health, Natural Resources and Applied Sciences

School of Natural and Applied Sciences

Department of Mathematics, Statistics and Actuarial Science

13 Jackson Kaujeua Street
Private Bag 13388
Windhoek
NAMIBIA

T: +264 61 207 2913
E: msas@nust.na
W: www.nust.na

QUALIFICATION : **BACHELOR OF SCIENCE IN APPLIED MATHEMATICS AND STATISTICS**	
QUALIFICATION CODE: **07BSAM**	LEVEL: **6**
COURSE: **REGRESSION ANALYSIS AND ANOVA**	COURSE CODE: **RAA602S**
DATE: **NOVEMBER 2023**	SESSION: **1**
DURATION: **3 HOURS**	MARKS: **100**

FIRST OPPORTUNITY: QUESTION PAPER

EXAMINER: Simon Pombili Kashihalwa

MODERATOR: Prof Rakesh Kumar

INSTRUCTIONS:

1. Answer all questions on the separate answer sheet.
2. Please write neatly and legibly.
3. Do not use the left side margin of the exam paper. This must be allowed for the examiner.
4. No books, notes and other additional aids are allowed.
5. Mark all answers clearly with their respective question numbers.

PERMISSIBLE MATERIALS:

1. Non-Programmable Calculator

ATTACHEMENTS

1. t- Table

This paper consists of 4 pages including this front page

QUESTION 1 [60]

1.1

For each of the following indicates the Statistical tools to be applied:

(a) Examining how the level of employee training impacts customer satisfaction ratings [2]

(b) Predict weight if we know an individual's height. [2]

(c) A researcher decides to study students' performance from a school over a period of time. [2]

(d) Measuring the relationship between two securities. [2]

(e) Predicting whether a patient has a particular disease or not. [2]

1.2

What is the effect of putting additional predictor variables in the model on each of the following?

a) R^2 [2]

b) $Adj\ R^2$ [2]

c) SSE [2]

d) Estimated standard deviation of the errors [2]

1.3 A biologist is comparing the intervals (m seconds) between the mating calls of a certain species of tree frog and the surrounding temperature (t ^{0}C). The following results were obtained.

t ^{0}C	8	13	14	15	15	20	25	30
m secs	6.5	4.5	6	5	4	3	2	1

a) Calculate the coefficient of correlation and interpret it [6]

b) Test the significance of the coefficient of correlation at 5% [8]

c) Find the equation of the regression line [4]

d) Use your regression line to estimate the time interval between mating calls when the surrounding temperature is 10 ^{0}C [2]

1.2 A multiple regression model of the following form is fitted to a data set.

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2)\ i.i.d.$$

Variables	Coefficients	Std. Error	t value	Pr(>|t|
Cons		0.196	8.438	3.57
x1	5.3036	2.5316		0.03
x2	4.0336		1.627	0.107
x3	-9.3153	2.4657		0.002
x4		2.2852	0.257	0.7973

a) Find the missing values in the output. [5]

b) Write down the estimated regression model. [2]

c) Test the significance of x4. [5]

d) Interpret the coefficients of all variables. [5]

e) Test the significance of x3. [5]

Question 2[40]

2.1 Study the ANOVA output and answer the questions that follow

Source of Variation	SS	df	MS	F	p-value	F crit
Between Groups	86049.55	.....	43024.78			
Within Groups	10254	......	1709	......	0.001	5.14
Total	96303.55	.....				

a) Mathematically state the two assumption of ANOVA (No mark for plain English). [4]

b) Formulate the hypothesis for treatment means and make a decision. [4]

c) Find the sample size used in this experiment. [1]

d) Find the total df, df for Between group and Within group. [3]

e) Calculate the test statistics. [3]

f) Write down the general format of the effects and Means model. [4]

2.2

Study the output below based on political election campaign, where the dependent variable is won a seat or not regressed by incumbency (0=challenger, 1=incumbent), spending measures in N$ and spending_total*c (interaction of spending and incumbency)

wonseat	Coef	Std.Err.	P>|Z|	95% CI
Incumb	3.200883	0.8391721	0.00	....
spend_total	0.0001604	0.0000232	0.00	0.001149 ; 0.0002058
spend_total*c	-.0000649	0.000428	0.130	....
_cons	-3.901699	0.429417	0.00	-4.743341 ; -3.060057

a) Explain why OLS is unsuitable for binary dependent variable. [2]

b) Calculate the 95% CI for incumb (β_1) and the 95%CI for OR. [8]

c) If a challenger spent N$10 000 more, what will be his or her odds of winning. [2]

d) Write down the estimated probability of winning a seat for the challenger. [3]

e) Calculate the 95% CI for spend_total*c(β_3) and comment on the significance of spend_total*c, using 95% CI. [6]

END

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659