



<b>QUALIFICATION : BACHELOR OF SCIENCE IN APPLIED MATHEMATICS AND STATISTICS</b>	
<b>QUALIFICATION CODE: 07BSAM</b>	<b>LEVEL: 6</b>
<b>COURSE:REGRESSION ANALYSIS AND ANOVA</b>	<b>COURSE CODE: RAA602S</b>
<b>DATE: JANUARY 2024</b>	<b>SESSION: 1</b>
<b>DURATION: 3 HOURS</b>	<b>MARKS: 100</b>

**SECOND OPPORTUNITY / SUPPLEMENTARY: EXAMINATION QUESTION PAPER**

**EXAMINER:** Mr 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. F-Table

**This paper consists of 4 pages including this front page**

**QUESTION 1 [46]**

1.1. For each of the following models state whether its parameters can be estimated using standard linear regression techniques. If linear regression can be used, what are the independent and dependent variables that should be used?

a)  $Y_i = \beta_0 + e^{\beta_1 x_i + e}$  [1]

b)  $Y_i = \frac{1}{\beta_0 + \beta_1 x_i + \epsilon_i}$  [3]

c)  $Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$  where  $\beta_2$  is known to be 5. [3]

1.2 A researcher is researching whether or not birds of prey exposed to pollutants lay eggs with thinner shells. He collects a random sample of egg shells from each of 6 different nests and tests for pollutant level, p and measures the thinning of the shell, t. the results are shown in the table below

p	3	8	30	25	15	12
t	1	3	9	10	5	6

- a) Explain why linear regression model may be appropriate to describe the relationship between p and t. [2]
- b) Calculate the coefficient of correlation. [5]
- c) Find the equation of the regression line. [8]
- d) The scientist reviews similar studies and finds that pollutant levels above 16 are likely to result in the death of a chick soon after hatching. Estimate the minimum thinning of the shell that is likely to results in the death of a chick. [2]

1.3 Here is some SPSS output from the multiple linear regression with independent variables slope, porosity, and cover.

Dependent variable: Soillost						
Source	DF	Sum of squares	Mean square	F-Value	Pr>F	
Model	3	680.68178	.....			
Error	.....	.....	.....	.....	<.0001	
Total	10	696.58545				

  

Parameter Estimates					
Variable	DF	Parameter	Standard error	t-value	Pr> t
Intercept	1	-1.59534	.....	-0.08841628	0.932
Slope	1	76.45678	44.29509	.....	0.128
Porosity	1	1.57585	.....	2.154979077	0.0681
Cover	1	-23.77054	.....	-1.78108244	0.1181

- a) Complete the ANOVA table. [5]
- b) Compute coefficient of determination. [2]

- c) Compute the adjusted coefficient of determination. [2]
- d) Compute the missing standard errors and the t-values. [4]
- e) Using matrix form, state the fitted multiple linear regression model. [5]
- f) Use ANOVA to test the significance of the regression coefficients. [4]

**Question 2 [54]**

**2.1**

Six different machines are being compared for use in manufacturing jackpots ram. The machines are being compared with respect to durability and performance of the ram. A random sample of 4 rams from each machine is used to determine whether the mean durability and performance varies from machine to machine.

Machine					
1	2	3	4	5	6
17.5	16.4	14.3	14.6	17.5	11.9
16.9	19.2	13.5	16.7	19.2	15.1
15.8	17.7	11.8	20.8	16.5	14.3
18.6	15.4	10.5	18.9	20.5	16.8

Perform the analysis of variance at 0.05 level of significance and indicate whether or not the mean durability and performance differ significantly for the six machines. [18]

**2.2**

Assume that, in Windhoek between 2000 and 2008, the growth curve of internet subscribers is of the form  $Y = ae^{bx}$  and that the following data were observed.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008
# of internet subscribers	498	872	1527	2672	4677	8186	14325	25069	43871

- a) Find the regression coefficients  $a$  and  $b$  . Hint use the sequential method to code the variable time, start with  $x = 1$  for 2000. [21]
- b) Predict the number of internet subscribers after four years. [4]

### 2.3

The ministry of safety and security carried out a study to identify factors associated with an individual owning a gun or not, factors such as Sex, Age, Race, education were assess and the model output presented below.

Variables in the equation							OR 95% CI
Variable	B	S.E	Wald	DF	sig	Odds Ratio	
Male Female*	-0.78	0.124	39.624	1	0.000		
Age	0.02	0.004	32.65	1	0.000		
White Black*	1.618	0.197	67.534	1	0.000		
Educat	-0.23	0.02	1.37	1	0.242		
Constant	-2.246	0.363	38.224	1	0.000		

\*Reference variable

- a) Construct the 95% confidence interval for each parameter. [5]
- b) Estimate the odds for each parameter. [5]
- c) Write down the estimated model. [1]

END

Critical Values of the  $F$ -Distribution:  $\alpha = 0.05$ 

Denom. d.f.	Numerator Degrees of Freedom									
	1	2	3	4	5	6	7	8	9	10
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.737	4.531	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.841	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.491	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.533	2.421	2.334	2.266	2.211	2.165
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255	2.199	2.153
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244	2.189	2.142
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235	2.179	2.133
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225	2.170	2.123
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217	2.161	2.114
36	4.113	3.259	2.866	2.634	2.477	2.364	2.277	2.209	2.153	2.106
37	4.105	3.252	2.859	2.626	2.470	2.358	2.270	2.201	2.145	2.098