

NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF APPLIED SCIENCES

SMA 4105

DEPARTMENT OF APPLIED MATHEMATICS

EXPERIMENTAL DESIGN & MULTIPLE REGRESSION

SUPPLEMENTARY EXAMINATION

JULY 2005

Time: 3 HOURS

Answer **FOUR** questions: Question 1 in **SECTION A** (28 marks) and **THREE** from **SECTION B** (24 marks each). Where a question contains subdivisions, the mark value of each subdivision is indicated in brackets.

SECTION A

1. An experiment was carried out to investigate the effect of changing time of exposure to heat (x) on bacteria present in a food product. Data collected show the average number of surviving bacteria (y) and minutes of exposure to 150°C heat (x) for 12 different samples as follows:

y	175	108	95	82	71	50	49	31	28	17	16	11
x	2	2	4	4	6	6	8	8	10	10	12	12

- (i) Plot a scatter diagram of the data on graph paper. Fit a straight-line model to the data and draw your fitted line on your scatter diagram. [6 marks]
- (ii) Test for the significance of the slope through the analysis of variance approach. [6 marks]
- (iii) Compute the PURE ERROR and LACK-OF-FIT sums of squares and hence carry out the Lack-of-fit test. Is your conclusion here in agreement with your plots in (i) above? [8 marks]

(b) In multiple linear regression the least squares estimator β of the regression coefficient is given by $\hat{\beta} = (X'X)^{-1}X'Y$.

(i) Show that $\hat{\beta}$ is an estimator of β . [4 marks]

(ii) Show that $\text{Var}(\hat{\beta}) = \sigma^2(X'X)^{-1}$. [4 marks]

SECTION B

2. A bacteriologist is interested in the effect of two different culture mediums and two different times on the growth of a particular virus. She performs six replicates of a 2^2 factorial experiment, making the runs in random order. The following results are obtained.

		Culture Medium			
Time		1		2	
12 hours		21	22	25	26
		23	28	24	25
		20	26	29	27
18 hours		37	39	31	34
		38	38	29	33
		35	36	30	35

- (i) Make a table of treatment combinations in standard order and their totals. [4 marks]
- (ii) Compute the sums of squares for both main effects and two factor interaction by the contrast method. [12 marks]
- (iii) Construct an ANOVA table and test for the significance of main effects and interaction. [8 marks]

3. Suppose that a scientist takes experimental data on the radius of a propellant grain (y) as a function of powder temperature (x_1), extrusion rate (x_2) and dye temp (x_3). The following data were obtained in one such experiment.

Grain radius (y)	Powder temperature (x_1)	Extrusion rate (x_2)	Die temperature (x_3)
82	150	12	220
93	190	12	220
114	150	24	220
124	150	12	250
111	190	24	220
129	190	12	250
157	150	24	250
164	190	24	250

- (a) Fit a multiple linear regression model to the data. [10 marks]
- (b) Partition the regression sum of squares into three single degree of freedom components attributable to x_1, x_2 and x_3 . Construct an analysis of variance table indicating significance on each regressor variable. Draw your conclusion. [14 marks]

4. An engineer is studying the effect of cutting speed on the rate of metal removal in a machining operation. However the rate of metal removal is also related to the hardness of the test specimen. Five observations are taken at each cutting speed. The amount of metal removed (Y) and the hardness of the specimen (X) are shown in the following table.

Cutting Speed					
1000		1200		1400	
Y	X	Y	X	Y	X
68	120	112	165	118	175
90	140	94	140	82	132
98	150	65	120	73	124
77	125	74	125	92	141
88	136	85	133	80	130
421	671	430	683	445	702

$$\begin{aligned} \sum Y &= 1296 & \sum Y^2 &= 115148 & \sum XY &= 180946 \\ \sum X &= 2056 & \sum X^2 &= 285366 & & \end{aligned}$$

NB The absence of subscripts indicates the sum of all X's, Y's and XY products.

- (a) Ignore the covariate X and test for the significant differences in the amount of the metal removed (Y) at the three cutting speeds by the Analysis of Variance (ANOVA). [8 marks]
- (b) Using the covariate X, Perform the analysis of covariance (ANCOVA) on the data. [12 marks]
- (c) Compare the results obtained in (a) with those obtained in (b). [4 marks]
5. A nickel titanium alloy is used to make components for jet turbine aircraft engines. Cracking is a potentially serious problem in the final part as it can lead to non recoverable failure. A test is run at the parts producer to determine the effect of four variables on cracks. The four variables are pouring temperature (x_1) titanium content (x_2), heat treatment method (x_3) and amount of grain refiner used (x_4). A total of 30 observations of y were recorded at various levels of x_1, x_2, x_3 , and x_4 . All possible linear

regression models which included a constant, were examined and the residual sums of squares RSS (p) are as follows.

Regression Included	RSS	p	C_p	S_p
Constant	2 715,7635			
x_1	1 265, 6867			
x_2	906,3363			
x_3	1 939,4005			
x_4	883,8669			
x_1, x_2	57,9045			
x_1, x_3	1 227,0721			
x_1, x_4	74,7621			
x_2, x_3	415,4427			
x_2, x_4	868,8801			
x_3, x_4	175,7380			
x_1, x_2, x_3	48,1106			
x_1, x_2, x_4	47,9727			
x_1, x_3, x_4	50,8361			
x_2, x_3, x_4	73,8145			
x_1, x_2, x_3, x_4	47,8636			

Complete the table and use the following methods to select the "best" linear regression model.

- (a) The c_p and s_p statistics. [8 marks]
- (b) Forward selection. [8 marks]
- (c) Backward elimination. [8 marks]

Show your working clearly and logically.

6. a) Construct a 2^2 design with ABC confounded in the first 2 replicates and BC confounded in the third. Outline the analysis of variance and comment on the information obtained. [10 marks]
- (b) Discuss clearly what is meant by each of the following giving an appropriate example for each.
- (i) Complete confounding [7 marks]
 - (ii) Aliasing. [7 marks]

END OF QUESTION PAPER