

SMA2215

NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF APPLIED MATHEMATICS

B.Sc. PART II HONOURS EXAMINATIONS 2004

SMA 2215 – INTRODUCTION TO APPLIED STATISTICS

JUNE 2004

3 HOURS (100 Marks)

This paper has 5 pages

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.

Candidates are expected to spend not more than one hour on Question 1. Calculators may be used. Statistical Tables and graph paper are provided, however, Statistical Tables should not be marked or taken out of the examination room. **GOOD LUCK!**

SECTION A (COMPULSORY)

1. (a) The probability that a patient recovers from a rare blood disease is 0.3. If 20 people are known to have contracted this disease, what is the probability that
 - i) at least 10 survive **(3 marks)**
 - (ii) from 3 to 8 survive **(3 marks)**
 - (iii) exactly 5 survive **(2 marks)**

- (b) As part of an air pollution survey, an inspector decides to examine the exhaust of six of a company's 24 trucks. If four of the company's trucks emit excessive amounts of pollutants, what is the probability that none of the four will be included in the inspector's sample? **(4 marks)**

- (c) Records show that the probability is 0.00005 that a car will have a flat tyre while crossing a certain bridge. What is the probability that among 10 000 cars that cross the bridge
 - (i) exactly two will have a flat tyre? **(2 marks)**
 - (ii) at most two will have a flat tyre? **(2 marks)**

- (d) The mean level of prothrombin in the normal population is known to be approximately 20 mg/100ml of plasma. A sample of 625 patients showing a vitamin K deficiency has a mean prothrombin level of 18.50 mg/100ml. The sample standard deviation is 4 mg/100ml. Do patients with vitamin K deficiency have a significantly lower prothrombin level than that of the general population? **(5 marks)**
- (e) An electrical firm manufactures light bulbs that have a length of life that is normally distributed with mean equal to 800 hours and standard deviation of 40 hours.
- (i) Find the probability that a bulb burns between 778 and 834 hours. **(3 marks)**
- (ii) Find the probability that a random sample of 16 bulbs drawn from this normal population will have an average life of less than 775 hours. **(4 marks)**

SECTION B (Answer any THREE questions)

2. (a) A medical doctor was interested in testing the effectiveness of a new drug product to cure acne. A random sample was chosen consisting of 180 people with approximately the same degree of acne. The people were then randomly divided into two groups of 90 each. Those in the first group were given the drug product while those in the second group received no treatment and hence served as a **control** group. After a predetermined length of time, each of the 180 people were examined. The acne condition of each person was classified as either **worse, same, or better**.

CONDITION	THERAPY		Total
	Drug treated	Control	
Worse	23		41
Same		52	
Better	43		
Total		90	180

- (i) Complete the table. **(2 marks)**
- (ii) Carry out an appropriate test at the 1% level of significance to test whether there is any association between condition of acne and therapy. **(5 marks)**
- (b) Two protein diets were allocated randomly to male chicks of the same age and breed. The weight (in grams) of chicks after four weeks are as follows:

DIET	WEIGHT (in gm)					
	High protein	150	136	144	168	152
Low protein	121	109	115	106	99	

- (i) Carry out a two-sample (pooled) t-test at the 5% level of significance to test whether there is a significant difference between the two protein diets. **(7 marks)**
- (ii) Carry out a one-way analysis of variance (ANOVA) to test whether there is a significant difference between the two protein diets at the 5% level of significance. **(8 marks)**
- (iii) Explain how the test in (i) is related to the test in (ii). **(2 marks)**
3. (a) A researcher believes that she has designed a keyboard that is more comfortable to use than a standard keyboard. In order to help decide if this is the case, typing speeds were taken for eight different people on each keyboard. The lengths of time, in minutes, for each of the people to type a pre-selected manuscript were recorded and are shown below:

Person	Original Keyboard	New Keyboard
1	15	12
2	9	8
3	17	15
4	10	8
5	9	5
6	4	4
7	30	25
8	29	21

State your assumptions and hence test at the 5% level of significance whether there is a significant difference between the two keyboards with respect to the mean time taken in typing the manuscript. **(9 marks)**

- (b) You are now told that sixteen people were chosen for the experiment. Eight of these were randomly assigned to use the original keyboard and the other eight to use the new keyboard in typing a pre-selected manuscript such that each pair of the observation in (a) above was recorded from two different people. State your assumptions and hence test at the 5% level of significance whether there is a significant difference between the two keyboards with respect to the mean time taken in typing the manuscript. **(10 marks)**
- (c) What are the advantages and disadvantages of the test in (a) over that in (b)? **(5 marks)**

4. The following are the cholesterol contents (in milligrams per package), which four laboratories obtained for packages of three very similar diet foods:

LABORATORY	DIET FOOD		
	A	B	C
1	3.4	2.6	2.8
2	3.0	2.7	3.1
3	3.3	3.0	3.4
4	3.5	3.1	3.7

- (a) A food nutritionist at the University of Zimbabwe decides that there are no systematic differences among the four laboratories and she therefore ignores the effect of laboratory in her analysis to determine whether significant differences exist in cholesterol content of the three diet foods. Carry out this analysis at the 5% level of significance. **(10 marks)**
- (b) A food nutritionist at NUST, however, suspects that there may be systematic differences among the four laboratories and she, therefore, includes this source of variation in her analysis to determine whether there are any significant differences in cholesterol content of the three diet foods. Carry out an analysis at the 5% level of significance. **(12 marks)**
- (c) Of the two approaches, (a) and (b), which one is more precise and why? **(2 marks)**
5. A limnologist collected the following data for a lake under study. The variables are $x =$ depth of lake (in metres) and $y =$ temperature (in degrees C)

Depth, x	4	17	3	11	5	12	7	8	13	2	9	1	16	6
Temp, y	17	5	16	11	16	12	14	14	11	18	13	18	6	15

$$\sum_{i=1}^{14} x_i = 114$$

$$\sum_{i=1}^{14} y_i = 186$$

$$\sum_{i=1}^{14} x_i y_i = 1256$$

$$\sum_{i=1}^{14} x_i^2 = 1264$$

$$\sum_{i=1}^{14} y_i^2 = 2682$$

- (a) Draw a scatter diagram of y against x on graph paper. Fit the least squares straight line to these data and draw your fitted line on your scatter diagram. **(10 marks)**
- (b) Carry out an analysis of variance (ANOVA) to test at the 5% level of significance whether the slope is significantly different from zero. From your ANOVA table, compute the coefficient of determination, r^2 , and interpret it. **(8 marks)**
- (c) Predict the temperature at a depth of 10 metres. Compute the standard error of the predicted value and hence construct the 95% confidence interval of this prediction. **(6 marks)**

6. A medical doctor is working on a drug to help cure depression. The doctor is interested in comparing three different doses of two different drug products. A sample of 36 patients all with similar tendencies of depression was randomly divided into two equal groups. The first group received drug Q, with 6 people randomly assigned to each of the three doses (100 mg, 200 mg, and 300 mg). Similarly, the people in the second group received drug R, with 6 people randomly assigned to each of the three doses. The doctor examined each patient 3 days after taking the drug. A score was assigned to each of the patients based on their mood and the following results were obtained:

DRUG	DOSAGE (mg)					
	Dose 100		Dose 200		Dose 300	
Q	120	120	131	128	135	137
	119	124	133	128	134	129
	110	121	127	130	130	136
R	127	130	128	133	160	158
	135	131	140	139	156	158
	132	130	142	135	160	149

- (a) Construct a two-way table of totals. (2 marks)
- (b) Carry out a preliminary ANOVA to test for the significance of treatment combinations at the 5% level of significance. (8 marks)
- (c) Carry out a complete ANOVA to test for the significance of the main effects and interaction at the 5% level of significance. (10 marks)
- (d) Give an appropriate table of means and appropriate standard errors for comparing any two means. (4 marks)

-----*** END OF QUESTION PAPER ***-----