SMA2115

NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF APPLIED MATHEMATICS

BSc HONOURS PART II EXAMINATIONS 2003

SMA 2115 – INTRODUTION TO APPLIED STATISTICS

NOVEMBER/ DECEMBER 2003

3 HOURS (100 Marks) This paper has 4 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) In each case identify the appropriate probability distribution and compute the required probability:

   i) The probability of a female birth in a certain area is 0.45. What is the probability that a couple who have five children will have at least three girls? **(4 marks)**

   ii) In a certain hospital, the probability of a multiple birth is 0.01. Suppose that you take a random sample of 20 deliveries. What is the probability of no more than one multiple birth? **(4 marks)**

   (iii) A certain birth defect occurs with probability 0.0001, that is one of every 10 000 babies has this defect. If 5 000 babies are born at a particular hospital in a given year, what is the probability that there is at least one baby with the defect. **(4 marks)**

   (iv) Among the 16 applicants for a job, ten have university degrees. If three of the applicants are randomly chosen for interviews, what is the probability that all three have university degrees. **(4 marks)**

   (b) A certain kind of appliance requires repairs on the average once every two years Assuming that the times between repairs are exponentially distributed, what is the probability that such an appliance will work at least three years without requiring repairs? **(4 marks)**

(c) In a photographic process, the developing time of prints may be looked upon as a random variable having a normal distribution with a mean of 15.40 seconds and a standard deviation of 0.48 second. Find the probabilities that the time it takes to develop one of the prints will be:

(i) at least 16.00 seconds, **(1 mark)**

(ii) at most 14.20 seconds, **(1 mark)**

(iii) anywhere from 15.00 to 15.80 seconds. **(2 marks)**

(d) If X and Y are random variables with a joint probability distribution such that Var (X) = 2, Var(Y) = 4 and Cov(X,Y) = -2

(i) find Var(Z), where Z = 3X – 2Y +5 **(1 mark)**

(ii) find Cov(P,Q), where P = 3X + 4Y; Q = X – 2Y **(3 marks)**

SECTION B (Answer any THREE questions)

2 (a) A total of 210 emphysema patients entering a clinic over a one-year period were treated with one of two drugs (either the standard drug, A, or an experimental compound, B) for a period of one week. After this period, each patient's condition was rated as either greatly improved, improved, or no change. The sample data are shown below:

	PATIENT'S CONDITION		
THERAPY	**No change**	**Improved**	**Greatly Improved**
Standard Drug, A	20	35	45
Experimental Drug, B	15	45	50

Test at the 5% level of significance whether the patient's condition is depended on the therapy used. **(10 marks)**

(b) Three groups of six guinea pigs each were injected, respectively, with 0.5 milligram, 1.0 milligram, and 1.5 milligrams of a new tranquilizer, and the following are the minutes it took them to fall asleep:

TRANQUILIZER LEVEL	**MINUTES**					
0.5 mg, A	21	23	19	24	25	23
1.0 mg, B	19	21	20	18	22	20
1.5 mg, C	15	10	13	14	11	15

Test at the 0.05 level of significance whether the null hypothesis that differences in dosage have no effect can be rejected. **(14 marks)**

3. The following data pertain to the chlorine residual in a swimming pool at various times after it has been treated with chemicals:

Number of hours, x	2	4	6	8	10	12
Chlorine residual, y (ppm)	1.8	1.5	1.4	1.1	1.1	0.9

(a) Draw a scatter diagram of y against x. Fit a least squares line to the 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. Compute the coefficient of determination, r^2, and interpret it. **(8 marks)**

(c) Predict the chlorine residual in the pool 5 hours after it has been treated with chemicals. Compute the standard error of the predicted value and hence find the 95% confidence interval for the predicted value. **(6 marks)**

4. An experiment was set up to compare the effect of different soil pH and calcium additives on the increase in trunk diameters for orange trees. Annual applications of elemental sulphur, gypsum, soda ash, and other ingredients were applied to provide pH value levels of 4, 5, 6 and 7. Three levels of a calcium supplement (100, 200 and 300 kg per ha) were also applied. All factor-level combinations of these two factors were used in the experiment. At the end of a two-year period, three diameters were examined at each factor-level combination and the following data obtained:

	Calcium		
pH Value	100	200	300
4.0	5.2 5.9 6.3	7.4 7.0 7.6	6.3 6.7 6.1
5.0	7.1 7.4 7.5	7.4 7.3 7.1	7.3 7.5 7.2
6.0	7.6 7.2 7.4	7.6 7.5 7.8	7.2 7.3 7.0
7.0	7.2 7.5 7.2	7.4 7.0 6.9	6.8 6.6 6.4

(a) Construct a two-way table of totals. **(2 marks)**

(b) Carry out a preliminary ANOVA to test for the significant differences among the twelve treatment combinations at the 5% level of significance. **(8 marks)**

(c) Carry out a complete ANOVA to test for the significance of main effects and interaction at the 5% level of significance. **(10 marks)**

(d) Give an appropriate table of means and appropriate standard errors for the comparisons. **(4 marks)**

5. (a) An experiment was conducted to compare the mean number of tapeworms in the stomachs of sheep that had been treated for worms against the mean number in those that were untreated. A sample of 14 worm-infected lambs was randomly divided into two groups. Seven lambs were injected with the drug and the remaining seven were left untreated. After a six-month period, the lambs were slaughtered and the following worm counts were recorded:

Drug-treated	18	43	28	50	16	32	13
Untreated	40	54	26	63	21	37	39

Make the relevant assumptions and test the hypothesis that there is no difference in the mean number of worms between treated and untreated lambs at the 5% level of significance. **(10 marks)**

(b) Suppose that you are now told that the data in (a) above were obtained from a set of seven lamb twins in which one member of a pair of twins was injected with the drug and the other left untreated.

Twin Pair	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Drug-treated	18	43	28	50	16	32	13
Untreated	40	54	26	63	21	37	39

Make relevant assumptions and test the hypothesis that there is no significant difference in the mean number of worms between treated and untreated lambs at the 5% level of significance. **(10 marks)**

(c) What are the advantages and disadvantages of the approach in (b) over that in (a)? **(4 marks)**

6. List the ADVANTAGES and DISADVANTAGES of each of the following experimental designs:

(a) The Completely Randomised Design. **(6 marks)**

(b) The Randomised Complete Block Design. **(6 marks)**

(c) The Latin Square Design. **(6 marks)**

(d) The Factorial Experiment. **(6 marks)**

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