

2003

ONLY candidates doing the course Statistics 1 and 2 and one unit chosen from Mathematics 1 (Section D), Numerical Analysis 1 (Section F) and Mechanics 1 (Section G) should attempt this Section.

Answer all the questions.

Answer these questions in a separate answer book, showing clearly the section chosen.

A1. A taxi was involved in a hit-and-run accident at night in the only city on an island. Only two taxi companies, the Green and the Yellow, trade in the city. The police investigation revealed that of the taxis in the city 85% are Green and 15% are Yellow and that there was only one witness to the accident.

Reliability testing of the witness revealed that, under conditions similar to those prevailing on the night of the accident, she correctly identified taxis from each of the companies 80% of the time and failed 20% of the time.

Using the above information, calculate the probability that the taxi involved in the accident was actually Yellow given that the witness claimed that it was Yellow.

5

A2. The MacBurger fast food restaurant chain has determined that, once outline planning permission has been granted, the time taken to obtain full planning permission for a new outlet has mean 180 days with standard deviation 12 days. The construction time for an outlet has mean 20 days and standard deviation 5 days.

On the assumption that the two phases, planning time and construction time, are independent, obtain the mean and standard deviation of the total time from the granting of outline planning permission to completion of construction.

3

On the day that outline planning permission for an outlet has been granted, the Marketing Department wishes to be able to display posters in the location of the site announcing "Your new MacBurger opening within n days!!!"

Stating any assumption required, calculate an integer value n such that there is at least 99.9% probability that the outlet will open within the specified number of days.

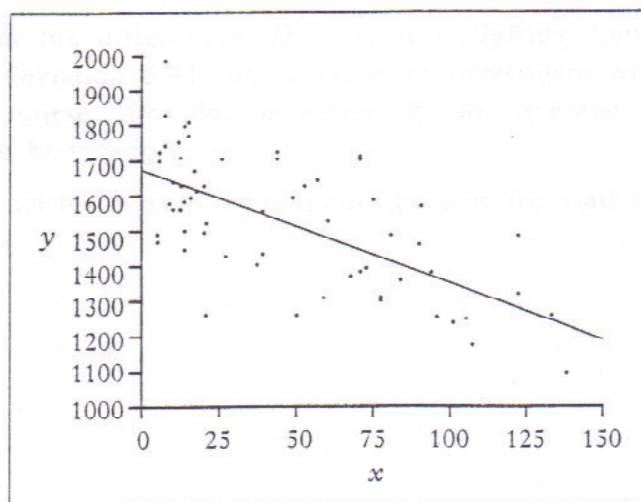
3

- A3. The number of pupils in each of the classes in the top three years in Newtown Secondary School is given below.

Year	Class	Number
6	6a	34
	6b	32
5	5a	28
	5b	30
4	4a	32
	4b	24
	4c	

- (a) Describe how you could use this data to enable you to select a systematic sample of 10% of senior pupils. 3
- (b) State an advantage and a disadvantage of systematic sampling. 2
- A4. A major construction company found that 75% of building projects which it undertakes are completed on schedule. Using a suitable approximation, calculate the probability that, out of 100 randomly selected projects, at most 70 will be completed on schedule. 6

- A5. An environmental scientist obtained annual mortality figures for males aged 55–64 averaged over the years 1958–1964 (y), and the calcium concentration (in parts per million) in drinking water (x) for sixty-one large towns in England and Wales. A scatter plot of the data showing the least squares regression line of y on x is given below.



The scientist calculated the sample correlation coefficient, r , to be -0.655 .

- (a) Test the null hypothesis that the population correlation coefficient is zero. 4
- (b) Interpret the result of the test in terms of the relationship between annual mortality and calcium concentration. 1

[Turn over

- A6. The number, X , of "crash calls" per month to the accident and emergency unit at a major hospital has the Poisson distribution with mean 10. Obtain the smallest integer a such that $P(10 - a < X < 10 + a) \geq 0.99$.

4

- A7. A financial institution gave a group of employees a test before and after a refresher course on tax legislation. The scores obtained are tabulated below.

<i>Employee</i>	<i>Score</i>	
	<i>Before</i>	<i>After</i>
1	48	56
2	87	89
3	82	79
4	44	53
5	56	58
6	71	66
7	60	64
8	66	80
9	84	87
10	48	53
11	63	71
12	48	47

- (a) Given that the differences, $D = \text{After} - \text{Before}$, have mean 3.83 and standard deviation 5.41, use a t -test to investigate whether or not the training course provides evidence of an average improvement in employees' knowledge.
- (b) State another test which it would have been appropriate to use.

5

1

- A8. A transportation company determined that the time for a bus journey, on a particular route during peak traffic periods, had mean 28 minutes with standard deviation 3 minutes. Following the introduction of a new traffic management system on the route the times for a random sample of journeys were as follows.

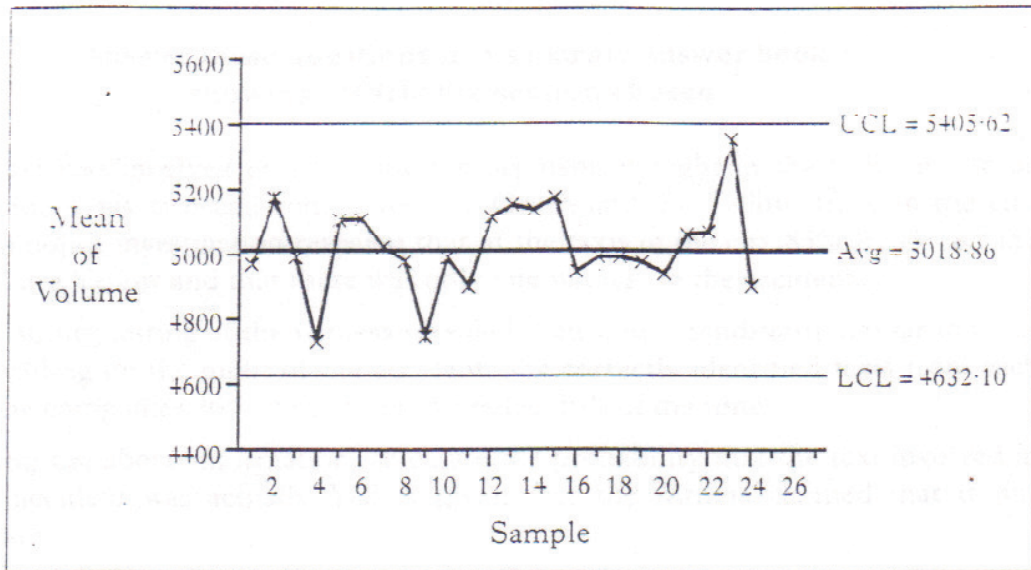
26	21	24	28	28	26	24	24
----	----	----	----	----	----	----	----

It can be assumed that the standard deviation of journey time has not changed.

- (a) Stating any further assumption required, use a z -test at the 1% level to determine whether or not there is evidence of a change in the mean time for the journey. 6
- (b) Calculate the p -value for the test and explain how it confirms your previous answer. 3
- (c) Explain how further confirmation of your answer is provided by the information that the 99% confidence interval for the mean journey time, calculated from the data, is (22.4, 27.9). 1
- A9. A physician extracted records at random on 100 stroke patients from a large database and found that 78 of them had suffered ischaemic strokes while the remainder had suffered haemorrhagic strokes.
- (a) Calculate an approximate 95% confidence interval for the proportion of ischaemic strokes in the population to which the patients belonged. 2
- (b) In another country the corresponding population proportion of ischaemic strokes is 0.65. State, with justification, whether or not there is evidence of differing proportions of ischaemic strokes amongst stroke patients in the two countries. 2
- Of the patients with ischaemic strokes 37 died as a direct result of the stroke while the corresponding figure for haemorrhagic strokes was 15.
- (c) Construct a contingency table from the data and carry out a test to determine whether or not the data provide evidence that survival rate is dependent on type of stroke. 7

[Turn over

A10. In order to monitor the volume of solder joints on computer memory modules, random samples of five joints were taken at regular intervals and the mean volume of each sample was displayed on a Shewhart control chart with three sigma limits as shown below.



- (a) Given that an estimate of the process standard deviation is 288.3, confirm the values for the control limits shown on the chart. Explain how the chart provides no evidence of special cause variation. 4
- (b) It is recommended that solder joint volume should lie between 4500 and 5500. Assuming that solder joint volume is normally distributed, estimate the proportion of joints which will meet the recommendation. 4
- (c) State two things which could be done with the process to ensure that the proportion meeting recommendations increased. 2

[END OF SECTION A]

All candidates who have attempted Section A (Statistics 1 and 2) should now attempt ONE of the following

Section D (Mathematics 1) on Page fifteen

Section F (Numerical Analysis 1) on Pages eighteen and nineteen

Section G (Mechanics 1) on Pages twenty and twenty-one.