Marks

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. Coloured cards are placed in three boxes as follows.

	Red	*White*	*Blue*
Box 1	3	4	3
Box 2	4	1	3
Box 3	2	3	5

A box is selected at random and a card drawn at random from it is found to be red.

Find the probability that Box 1 was selected.   5

A2. Scientists observed that the mean number of alpha particles emitted from a radioactive source during time intervals of 7·5 seconds was 4. Assuming that the situation can be modelled by a Poisson distribution, obtain the probability that during a randomly selected interval of 7·5 seconds the number of alpha particles emitted was

(*a*) zero,   1

(*b*) four,   2

(*c*) greater than four.   2

What is the standard deviation of the number of alpha particles emitted during intervals of duration 7·5 seconds?   1

A3. In order to carry out a study of farms in a region, the farms were classified as large, medium and small. Local authority records revealed the following numbers of farms in each category.

Category of farm	*Number of farms in category*
Large	30
Medium	320
Small	150

Explain how you would set about taking a stratified random sample of 50 farms in the region, specifying the number of farms in each category which you would sample.   3

State **two** advantages of stratified random sampling over simple random sampling.   2

A4. (*a*) Write down an expression for the 95% confidence interval for a population proportion *p*.

(*b*) Show that the maximum value of $p(1 - p)$ is $0 \cdot 25$.

(*c*) Use these results to show that the maximum width of such a 95% confidence interval is approximately $0 \cdot 06$ when the sample size is 1000.

A5. A market researcher believes that in a population of users of laptop computers 30% prefer manufacturer A, 60% prefer manufacturer B with the remainder preferring other manufacturers. A simple random sample of 600 users selected from the population showed the following preferences.

Manufacturer	A	B	Other
Number preferring this manufacturer	192	342	66

Do these data provide evidence to cast doubt on the researcher's belief?

A6. In a population of healthy adult males, systolic blood pressure, X (mmHg), is normally distributed with mean 140 and standard deviation 15. For the 99·9% normal range of systolic blood pressure in the population, physicians use the interval (*a*, *b*), with mid-point 140, such that $P(a < X < b) = 0 \cdot 999$. Calculate the values of *a* and *b*.

A7. A machine used for the automatic filling of breakfast cereal packs is known to deliver amounts of cereal with weights which are distributed with a standard deviation of 3 grams. When set for a production run, with packs labelled as 500 grams weight, the cereal manufacturer uses 505 grams as a target for the mean weight of the amounts. This is to ensure compliance with statutory regulations concerning under-filling of packs. During set-up, a trial run was carried out and a random sample of packs was found to have the following content weights.

509·9 507·7 506·7 507·5 506·7 506·1 507·0 506·5 505·2 510·5

Stating any assumption required, test the null hypothesis that the machine is operating "on target" using a two-tail test.

Explain why it would not be in the manufacturer's interest to simply carry out a one-tail test with alternative hypothesis that the mean amount is less than 505 grams.

[Turn over

A8. There is constant probability p that a computer "chip" produced at a factory is defective. State the distribution of $\dot{X}$, the number of defective chips, in random samples of size n and show that the sample proportion of defective chips has mean p and variance pq/n.

4

The production manager estimates that, for the current manufacturing process, p is 0·20 and she wishes to monitor the defective rate using daily random samples of 80 chips.

Find the probability that, with the specified proportion, a sample contains between 9 and 23 defective chips (inclusive), justifying, and showing clearly, your method.

6

A9. In the nineteenth century the Scottish physicist J. D. Forbes investigated the feasibility of estimating altitude in the mountains by observing the temperature at which water boiled. The table below gives data he collected on atmospheric pressure, x (inches Hg), and the corresponding boiling point of water, y (°F), from 17 locations in Scotland and the Alps. The table also gives the residuals from fitting the least squares regression of $y = 155\cdot3 + 1\cdot902x$.

x	y	*Residual*
20·79	194·5	−0·335
20·79	194·3	−0·535
22·40	197·9	0·004
22·67	198·4	−0·010
23·15	199·4	0·077
23·35	199·9	0·197
23·89	200·9	0·170
23·99	201·1	0·180
24·02	201·4	0·423
24·01	201·3	0·342
25·14	203·6	0·493
26·57	204·6	−1·227
28·49	209·5	0·022
27·76	208·6	0·510
29·04	210·7	0·176
29·88	211·9	−0·222
30·06	212·2	−0·264

(a) State the underlying model. Given that the sum of the squares of the residuals is 2·960, obtain an estimate of σ^2 in the model.

2

(b) Given that $S_{xx} = 145\cdot94$, test the null hypothesis that the slope parameter in the model is zero, stating any assumption required.
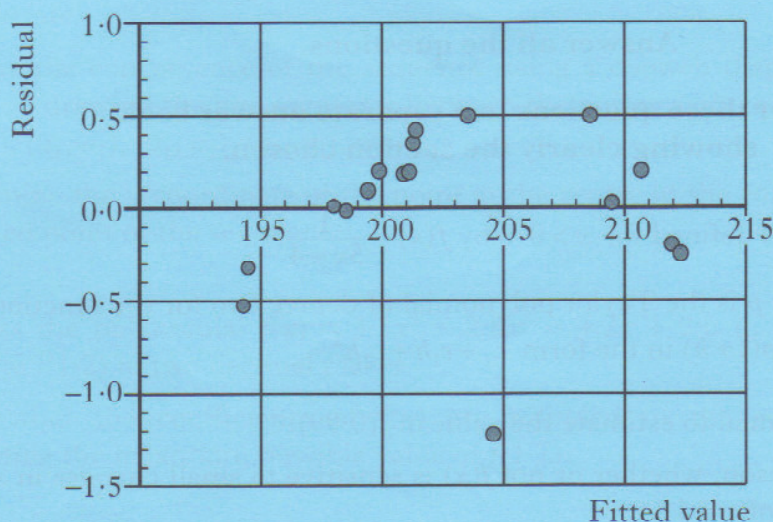
5

A9. **(continued)**

(c)   A plot of residuals against fitted values is shown below. Comment.   2

A10. In an experiment, designed to investigate the effect of caffeine on manual dexterity, twenty male students were trained to carry out a finger-tapping exercise.   Following the training ten of the students were given a drink containing a 200 mg dose of caffeine, while the remaining students were given a similar drink containing no caffeine.   Two hours later each student was observed carrying out the finger-tapping exercise and the following ordered data recorded.

Caffeine dose (mg)	Number of finger-taps per minute									
0	242	242	242	244	244	245	246	247	248	248
200	245	246	246	248	248	248	250	250	250	252

(a)   State **two** procedures which would constitute good practice during the running of the experiment.   2

(b)   Analyse the data using the Mann-Whitney test.   7

(c)   Describe briefly an alternative design for the experiment and state how you would analyse the resulting data.   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 fourteen*

Section F (Numerical Analysis 1) on *Pages seventeen* and *eighteen*

Section G (Mechanics 1) on *Pages nineteen* and *twenty*.