Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
TOTAL	

General Certificate of Education
Advanced Subsidiary Examination
June 2011

Statistics SS02

Unit Statistics 2

Thursday 26 May 2011 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions in the spaces provided.

1 A club introduces a weekly drum-and-bass night hosted by a DJ. The manager uses three DJs named Ed, Ja and Riz in rotation. For the first eleven weeks, the table shows the DJs used, the attendances and the values of a suitable moving average.

Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
DJ	**Ed**	**Ja**	**Riz**	**Ed**	**Ja**	**Riz**	**Ed**	**Ja**	**Riz**	**Ed**	**Ja**
Attendance	312	351	431	363	391	492	394	421	542	451	458
Moving average		364.7	381.7	395.0	415.3	425.7	435.7	452.3	471.3	483.7	

The graph opposite shows the attendances.

(a) (i) Add the moving averages to the graph opposite and draw a trend line. *(2 marks)*

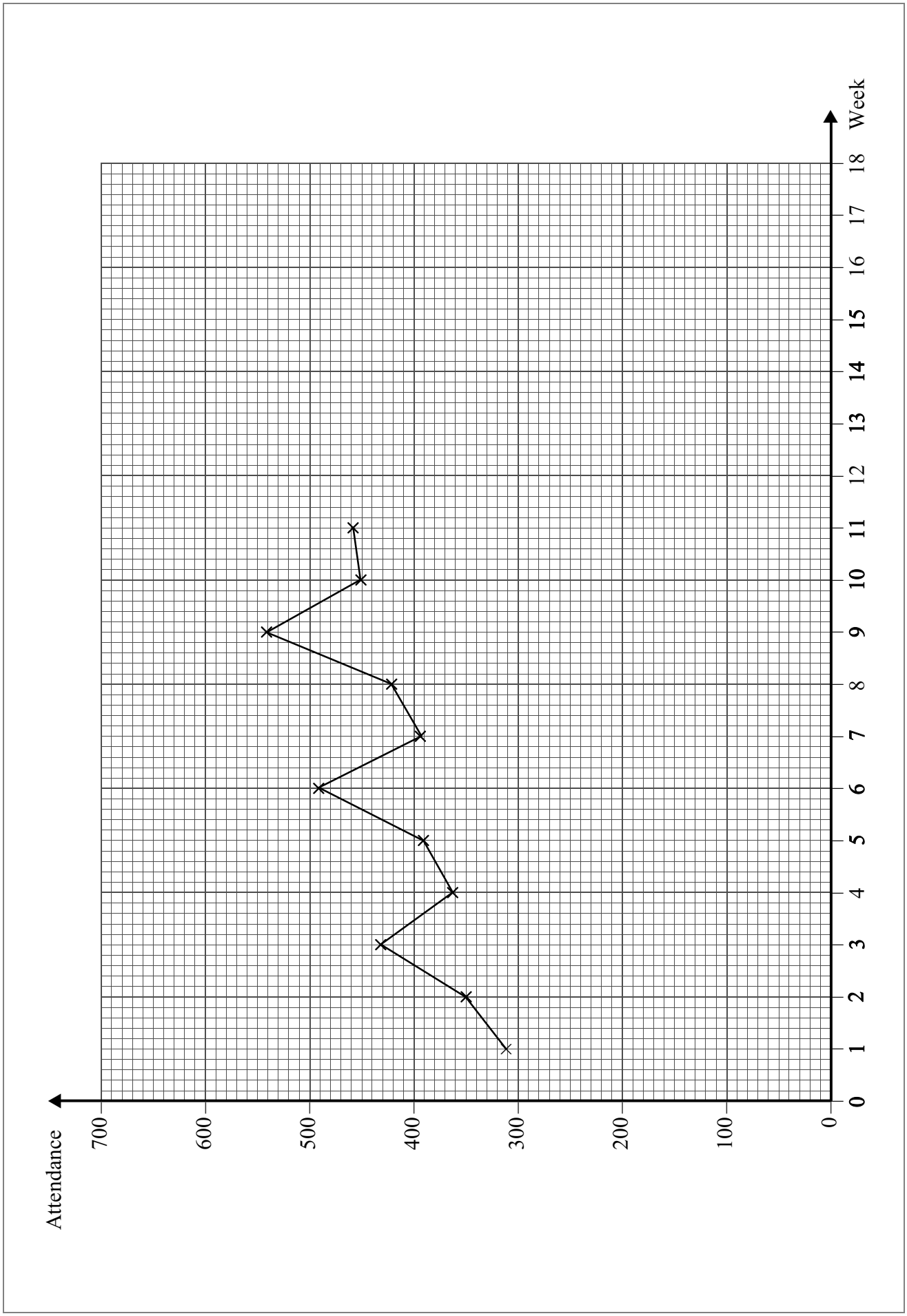
(ii) Estimate the 'seasonal' effect for Riz. *(3 marks)*

(iii) Forecast the attendance in week 15 when Riz is the DJ. *(3 marks)*

(b) For safety reasons, the attendance must be limited to 640. If the manager thinks this limit may be reached, extra security staff must be employed.

Advise the manager in which week it is likely that the extra security staff will first be needed if he continues to use the three DJs in the same rotation and current trends in attendance continue. Explain your answer. *(3 marks)*

QUESTION PART REFERENCE


Attendance
700
600
500
400
300
200
100
0
Week
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

QUESTION PART REFERENCE

QUESTION PART REFERENCE

2 In a pay-and-display car park, users are charged different amounts according to the lengths of time they wish to park their cars. The following table shows the distribution of the amounts, X pence, paid by users.

x	$\mathbf{P}(X = x)$
100	0.22
200	0.31
300	0.21
400	0.12
600	0.14

(a) (i) Show that the mean of X is 279.

(ii) Find the standard deviation of X. *(4 marks)*

(b) A small proportion of the cars in the car park belong to employees of the firm which operates the car park, who are allowed to park their cars with no charge. If these employees were included as users, state, giving a reason, whether the standard deviation would increase, stay the same or decrease. *(2 marks)*

(c) In fact, there is no charge for cars entering the car park after 6 pm. In the evening, it is used by a large number of people going to a nearby cinema. At 9 pm, a few cars which have been parked before 6 pm remain in the car park, but nearly all the cars in the car park have parked with no charge.

State, giving a reason, whether the standard deviation of the amounts paid to park the cars which are in the car park at 9 pm is greater than, the same as or less than the standard deviation of X. *(2 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

3 The table below shows the age distribution, in total and by sex, of the resident population of the United Kingdom as enumerated in each census from 1901 to 2001.

	United Kingdom population enumerated in Census — Thousands						
	1901	**1931**	**1951**	**1971**	**1981**	**1991**	**2001**
Persons: All ages	38 237	46 038	50 225	55 928	56 357	57 439	59 114
Under 1	938	712	773	899	730	790	663
1–4	3 443	2 818	3 553	3 654	2 726	3 077	2 819
5–9	4 106	3 897	3 689	4 684	3 677	3 657	3 735
10–14	3 934	3 746	3 310	4 232	4 470	3 485	3 890
15–19	3 826	3 989	3 175	3 862	4 735	3 719	3 678
20–29	6 982	7 865	7 154	7 968	8 113	9 138	7 499
30–44	7 493	9 717	11 125	9 797	10 956	12 125	13 405
45–59	4 639	7 979	9 558	10 202	9 540	9 500	11 168
60–64	1 067	1 897	2 422	3 222	2 935	2 888	2 884
65–74	1 278	2 461	3 689	4 764	5 195	5 067	4 947
75–84	470	844	1 555	2 159	2 677	3 119	3 296
85 and over	61	113	224	485	603	873	1 130
Males: All ages	18 492	22 060	24 118	27 167	27 412	27 909	28 832
Under 1	471	361	397	461	374	403	338
1–4	1 719	1 423	1 818	1 874	1 400	1 572	1 445
5–9	2 052	1 967	1 885	2 401	1 889	1 871	1 913
10–14	1 972	1 892	1 681	2 175	2 295	1 784	1 993
15–19	1 898	1 987	1 564	1 976	2 424	1 905	1 879
20–29	3 293	3 818	3 509	4 024	4 103	4 578	3 744
30–44	3 597	4 495	5 461	4 938	5 513	6 045	6 645
45–59	2 215	3 753	4 493	4 970	4 711	4 732	5 534
60–64	490	894	1 061	1 507	1 376	1 390	1 412
65–74	565	1 099	1 560	1 999	2 264	2 272	2 308
75–84	196	335	617	716	922	1 146	1 308
85 and over	23	36	70	126	141	212	312
Females: All ages	19 745	23 978	26 107	28 761	28 946	29 530	30 281
Under 1	466	351	376	437	356	387	324
1–4	1 724	1 397	1 735	1 779	1 327	1 505	1 375
5–9	2 054	1 930	1 804	2 283	1 788	1 786	1 822
10–14	1 962	1 854	1 629	2 057	2 175	1 701	1 897
15–19	1 928	2 002	1 611	1 887	2 311	1 815	1 799
20–29	3 690	4 047	3 644	3 945	4 009	4 560	3 755
30–44	3 895	5 222	5 663	4 859	5 442	6 080	6 760
45–59	2 424	4 226	5 065	5 231	4 829	4 769	5 634
60–64	577	1 003	1 361	1 715	1 559	1 498	1 473
65–74	713	1 361	2 127	2 765	2 931	2 795	2 640
75–84	274	509	937	1 443	1 756	1 972	1 987
85 and over	38	77	154	359	462	661	817

Source: *Annual Abstract of Statistics*, Office for National Statistics, 2008

(a) How many males aged '20–29' were enumerated in the 1951 census? *(2 marks)*

(b) The total of persons aged 'Under 1' enumerated in the 1901 census is not equal to the number of males aged 'Under 1' plus the number of females aged 'Under 1' enumerated in this census. Does this show that there must be an error in the data? Justify your answer. *(2 marks)*

(c) To what extent, if at all, does the table opposite provide evidence to support each of the following two common beliefs? Justify your answers.

(i) The probability of a baby being female is 0.5 . *(2 marks)*

(ii) On average, females live longer than males. *(2 marks)*

QUESTION PART REFERENCE

Question 3 continues on the next page

3 (d) (i) Complete **Figure 1** opposite, using data from the **2001** census, shown in the table on page **8**. *(3 marks)*

(ii) Illustrate the data in **Figure 1** by drawing a line diagram on **Figure 2** opposite. Add both a scale and a label to the vertical axis. *(3 marks)*

(iii) Comment on this age distribution. *(2 marks)*

(iv) Explain why a line diagram using the age groups shown in **the table on page 8** would be of little value. *(1 mark)*

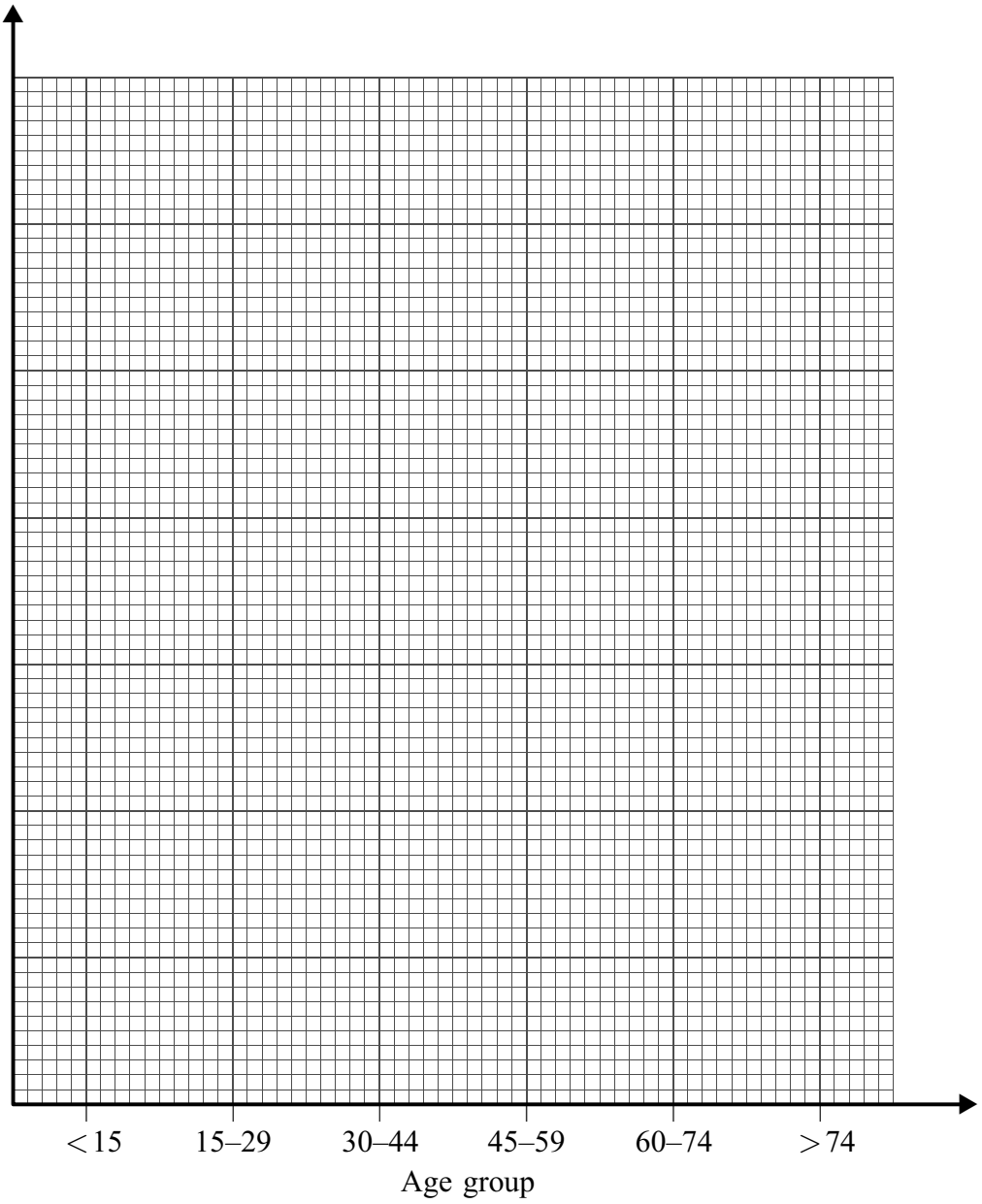
QUESTION PART REFERENCE

Figure 1

Males aged	Under 15	15–29	30–44	45–59	60–74	over 74
Thousands						

Figure 2

Males enumerated in 2001 Census
United Kingdom



4 (a) The number of reports of lost credit cards made to a bank's head office may be assumed to follow a Poisson distribution with mean 1.2 per hour. Find the probability that during a particular hour there will be:

(i) no reports;

(ii) exactly 1 report;

(iii) 5 or more reports. *(5 marks)*

(b) The bank's head office is open from 9 am to 6 pm on Mondays to Fridays: a total of 45 hours each week. A manager wishes to examine whether reports of lost credit cards arrive at random times during the week. He therefore records, for each hour of one week, the number of credit cards reported lost. The data are summarised in the table.

Number of reports	Number of hours
0	22
1	9
2	5
3	4
4	2
5	3

Find the mean and the variance of the number of reports of lost credit cards each hour. *(2 marks)*

(c) Explain why the Poisson distribution may not provide a good model for the number of reports of lost credit cards by:

(i) comparing your answer to part **(a)(i)** with the observed data in part **(b)**; *(2 marks)*

(ii) considering your answers to part **(b)**; *(1 mark)*

(iii) considering the context of this question. *(2 marks)*

(d) Explain to the manager whether or not the observed data support the view that reports of lost credit cards arrive at random times during the week. *(2 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

QUESTION PART REFERENCE

5 A company manufactures bath panels. The bath panels should be 700 mm deep, but a small amount of variability is acceptable. The depths are known to be normally distributed with standard deviation 2.1 mm.

(a) In order to check that the mean depth is 700 mm, Amir takes a random sample of 6 bath panels from the current production and measures their depths, in millimetres, with the following results.

701.2 698.2 704.4 699.4 695.5 698.9

Test whether the current mean is 700 mm, using the 5% significance level. *(8 marks)*

(b) Isabella, a manager, tells Amir that, in order to check whether the current mean is 700 mm, it is necessary to take a larger sample. Amir therefore takes a random sample of size 40 from the current production and finds that the mean depth is 701.34 mm.

Test whether the current mean is 700 mm, using the data from this second sample and the 5% significance level. *(5 marks)*

(c) It is proposed to carry out hypothesis tests at regular intervals to check that the mean remains at 700 mm.

Amir proposes that the tests be based on random samples of size 6, but Isabella favours random samples of size 40. Explain which, if either, sample size would lead to a smaller risk:

(i) of a Type I error;

(ii) of a Type II error. *(4 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

6 A school has 308 pupils in Year 8. Raina wishes to select a sample of 12 of these pupils to answer a questionnaire on eating habits.

(a) Describe how the table of random numbers in the booklet of formulae and statistical tables (Table 13) could be used to select a random sample of 12 of these 308 pupils. A list of the pupils' names is available. *(4 marks)*

(b) Emma numbers the pupils from 000 to 307. She then selects 3-digit random numbers from Table 13, divides each number by 308 and takes the remainder.

For example, she selects 637;
$637 \div 308 = 2$ remainder 21;
she selects pupil numbered 021.

She selects 12 pupils in this way, ignoring any repeats.

By considering the probability of selection of each of the pupils numbered 000 **and** 307, explain why her method would not lead to a random sample of pupils. *(2 marks)*

(c) John also numbers the pupils from 000 to 307. He uses his calculator to generate 12 random numbers between 0 and 1, each of which he then multiplies by 308 and rounds to the nearest whole number.

For example, he generates 0.292;
$0.292 \times 308 = 89.936$;
he rounds 89.936 to 90 and selects pupil numbered 090.

He selects 12 pupils in this way, ignoring any repeats and any numbers greater than 307.

Explain why the pupil numbered 000 would have less chance of being selected than the other pupils. *(2 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED