General Certificate of Education
January 2008
Advanced Subsidiary Examination

STATISTICS   **SS02**
Unit Statistics 2

AQA

ASSESSMENT and
QUALIFICATIONS
ALLIANCE

Friday 11 January 2008   9.00 am to 10.30 am

For this paper you must have:
- an 8-page answer book
- the blue AQA booklet of formulae and statistical tables
- an insert for use in Questions 1 and 2 (enclosed).

You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is SS02.
- Answer **all** questions.
- Show all necessary working; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.
- Fill in the boxes at the top of the insert.

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

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

SS02

Answer **all** questions.

1   [**Figure 1**, printed on the insert, is provided for use in this question.]

Emily works in an office. The box and whisker plot on **Figure 1** illustrates the number of minutes late that she arrived at the office on working days during 2006.

The following data are the number of minutes late that she arrived at the office for her 21 working days during January 2007. (Negative values indicate that she arrived early.)

15   37   22   8   23   44   16   22   46   2   −3   9   7   −5   4   −1   3   7   1   5   −2

(a)   For these data:

the lower quartile is 1.5;
the median is 7;
the upper quartile is 22;
there are no outliers.

Add a box and whisker plot for Emily's 21 working days during January 2007 to **Figure 1**.   *(3 marks)*

(b)   Make three comments comparing Emily's punctuality during January 2007 with that during 2006.   *(3 marks)*

(c)   The data above, for Emily's number of minutes late during January 2007, are presented in date order. Identify one important feature which is not illustrated by the box and whisker plot.   *(2 marks)*

2 [**Figure 2**, printed on the insert, is provided for use in this question.]

Ted is a qualified karate instructor and decides to hold classes in a sports centre on Monday, Wednesday and Friday evenings.

The following table shows the attendance for the first four weeks of these classes.

Week	1			2			3			4		
Day	**Mon**	**Wed**	**Fri**	**Mon**	**Wed**	**Fri**	**Mon**	**Wed**	**Fri**	**Mon**	**Wed**	**Fri**
Attendance	7	10	16	12	13	18	10	15	22	11	17	24

(a) Calculate values of a suitable moving average.   *(3 marks)*

(b) Plot your moving average values on **Figure 2** and draw a trend line by eye.   *(3 marks)*

(c) Use this trend line and your estimate of the 'Monday effect' to forecast the attendance on Monday of week 5.   *(5 marks)*

(d) Ted decides that unless average attendance is at least 19 in week 6 he will stop holding classes.

State, with justification, whether you think it is likely that Ted will stop holding classes.
                                                                                                              (2 marks)

Turn over for the next question

3 A bus to the city centre is scheduled to stop at Beech Road at 9.30 am on weekday mornings. The number of passengers getting on this bus at Beech Road may be modelled by a Poisson distribution with mean 8.5 .

  (a) Find the probability that on a particular weekday morning the number of passengers getting on this bus at Beech Road is:

    (i) 6 or fewer;

    (ii) more than 9. *(4 marks)*

  (b) The number of passengers getting **off** this bus at Beech Road on weekday mornings is independent of the number getting on and may be modelled by a Poisson distribution with mean 0.5 .

    Find the probability that on a particular weekday morning:

    (i) exactly 2 passengers will get off this bus at Beech Road;

    (ii) this bus will not need to stop at Beech Road as no passenger wishes to get on or off. *(4 marks)*

  (c) Serena, the bus driver, complains that the timetable does not allow her sufficient time to drive safely and to reach the city centre on schedule. The manager claims that there is enough time because it is unlikely that she will need to stop at every bus stop on the route.

    Without further calculation, comment on the claim that she is unlikely to need to stop at every bus stop. Assume that, for each bus stop on the route, the probability of not needing to stop is similar to that at Beech Road. *(2 marks)*

4 At the Berbekan Bakery a machine is used to measure out quantities of dough for making loaves. For each batch, the mean amount of dough per loaf is set using a control on the machine. The weights of the resulting loaves are normally distributed with standard deviation 24 grams.

The control is set to produce a batch of loaves with a mean weight of 1000 grams. The weights, in grams, of a random sample of nine loaves from this batch were

    998   996   936   1002   957   968   920   943   1011

Using the 5% significance level, examine whether the mean weight of this batch of loaves is 1000 grams. *(8 marks)*

5 A library has branches in Molcar and Garsden. The Molcar branch has 4 computer terminals available for the use of library members. Muttiah, a librarian, records the number of terminals, X, in use at random times during weekdays. He observes that X may be modelled by the following probability distribution:

x	$P(X = x)$
0	0.005
1	0.015
2	0.080
3	0.150
4	0.750

(a) (i) Show that the mean of X is 3.625.

(ii) Find the standard deviation of X. *(5 marks)*

(b) Muttiah also records the number of computer terminals, Y, in use at random times during weekdays in the library at Garsden. He observes that Y may be modelled by a Poisson distribution with mean 6.5.

(i) Find the standard deviation of Y. *(2 marks)*

(ii) Give a reason why, although Y may be adequately modelled by a Poisson distribution, it cannot follow a Poisson distribution exactly. *(1 mark)*

(iii) An extra terminal becomes available. Muttiah states that the mean number of computer terminals in use is smaller at Molcar than at Garsden and the standard deviation is also small. Hence he argues that there is little demand for an extra terminal at Molcar and therefore it should be installed at Garsden.

Comment on Muttiah's reasoning and on his conclusion. *(3 marks)*

Turn over for the next question

Turn over ▶

6 The following table refers to the sale of dwellings in London and the South East of England from 1994 to 2004.

Housing market: simple average house prices, mortgage advances and incomes of borrowers, by new/other dwellings, type of buyer and region

	Percentage of mortgages to first time buyers	New dwellings			Other dwellings		
		Average dwelling price	Average advance	Average recorded income of borrowers	Average dwelling price	Average advance	Average recorded income of borrowers
		£	£	£	£	£	£
London							
1994	61.7	75 200	56 054	24 960	87 563	63 775	29 278
1995	57.5	83 933	63 683	30 789	88 277	65 271	29 790
1996	50.4	99 292	71 189	32 232	93 321	68 667	31 308
1997	48.4	116 242	80 106	36 453	104 827	75 384	33 559
1998	55.9	125 079	91 041	38 734	114 166	81 023	35 193
1999	54.5	178 274	124 335	54 469	140 347	96 336	40 592
2000	51.0	210 655	139 193	57 532	161 145	109 963	43 538
2001	46.9	218 981	153 791	59 777	183 246	122 093	49 323
2002	38.6	239 360	162 777	61 710	205 120	137 359	52 360
2003	33.9	314 622	208 717	68 730	238 995	151 735	52 506
2004	35.2	310 613	217 970	73 495	268 686	172 801	55 315
South East							
1994	47.2	93 689	65 278	29 025	80 187	56 421	25 718
1995	44.0	105 474	73 733	34 242	79 340	57 083	25 756
1996	41.0	113 476	78 862	36 116	84 520	59 789	27 803
1997	37.9	121 488	83 388	37 301	91 782	64 993	29 587
1998	39.2	126 333	87 781	37 475	103 847	70 152	31 708
1999	38.0	165 225	105 988	45 442	117 575	78 147	34 274
2000	35.5	184 802	114 895	47 331	138 856	90 651	37 079
2001	33.6	196 079	126 583	51 887	154 733	96 651	40 758
2002	29.9	231 279	142 319	54 949	176 293	107 292	42 787
2003	23.0	276 318	169 893	58 613	212 517	126 376	45 583
2004	22.4	280 000	173 049	60 052	233 076	138 553	45 526

Source: *Housing Statistics*, Office of the Deputy Prime Minister, 2005

(a) For London during 1999, what percentage of mortgages were to first time buyers?

(1 mark)

(b) For new dwellings:

   (i) describe the trend in the average dwelling price in London from 1994 to 2004;

   (ii) compare the average dwelling price from 1994 to 2004 in the South East with that in London. *(5 marks)*

(c) For London, calculate the ratio of the average advance to the average recorded income of borrowers for:

   (i) new dwellings in 1994;

   (ii) other dwellings in 1994;

   (iii) new dwellings in 2004;

   (iv) other dwellings in 2004.

   Compare your ratios and briefly summarise your findings. *(4 marks)*

7 Illustrate, by means of a sketch, a time series which exhibits:

(a) random variation about a downward linear trend;

(b) short-term variation about an upward linear trend. *(4 marks)*

Turn over for the next question

Turn over ▶

8 A university employs 820 staff at a city centre site.

Natasha, an administrator in charge of car parking, wishes to survey members of staff as to their views on the present parking arrangements. A list of the 820 members of staff is available.

(a) Describe how Natasha could select a simple random sample of size 25 for the survey.
(4 marks)

(b) The university has a car park with 200 parking spaces. The spaces are numbered from 1 to 200. There are 300 members of staff with permits to park in this car park and each morning the spaces are allocated on a first come, first served basis. Permit holders arriving when the car park is full have to park elsewhere. There are 220 members of staff on a waiting list for a permit and the remaining 300 members of staff do not wish to have a permit.

  (i) It is suggested that, instead of a simple random sample, Natasha should take a stratified sample. Suggest two relevant factors that Natasha could use in the stratification of the staff.
(2 marks)

  (ii) A second suggestion is as follows:

  Select a systematic sample of 25 parking spaces. Natasha and her assistant will wait in the car park in the morning and when a car parks in one of the selected spaces the driver will be given a questionnaire and asked to complete it.

  (A) Describe how the systematic sample of 25 parking spaces could be selected.

  (B) State one advantage of this method of obtaining completed questionnaires.

  (C) State two sources of bias in this method of data collection.
(5 marks)

END OF QUESTIONS

Surname		Other Names	
Centre Number		Candidate Number	
Candidate Signature			

General Certificate of Education
January 2008
Advanced Subsidiary Examination

STATISTICS **SS02**
Unit Statistics 2

AQA
ASSESSMENT and
QUALIFICATIONS
ALLIANCE

Insert

Insert for use in **Questions 1 and 2**.

Fill in the boxes at the top of this page.

Fasten this insert securely to your answer book.

Turn over for Figure 1

Turn over ▶

Figure 1 (for use in Question 1)

Emily's Punctuality

Figure 2 (for use in Question 2)

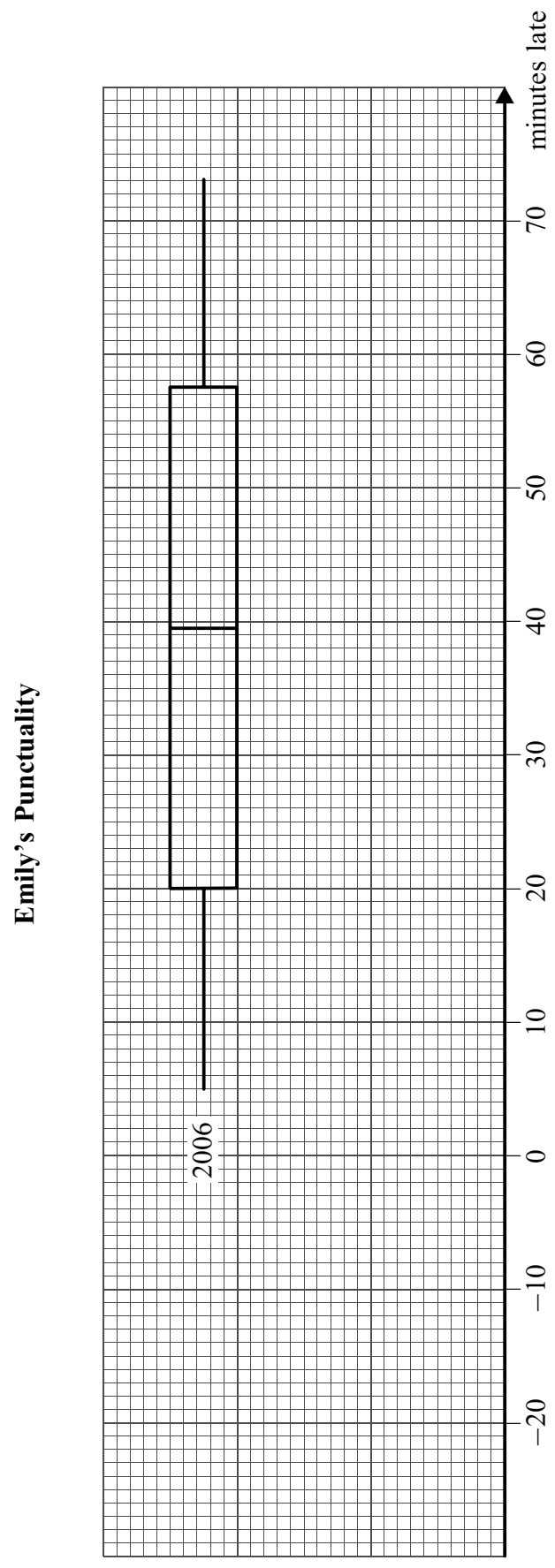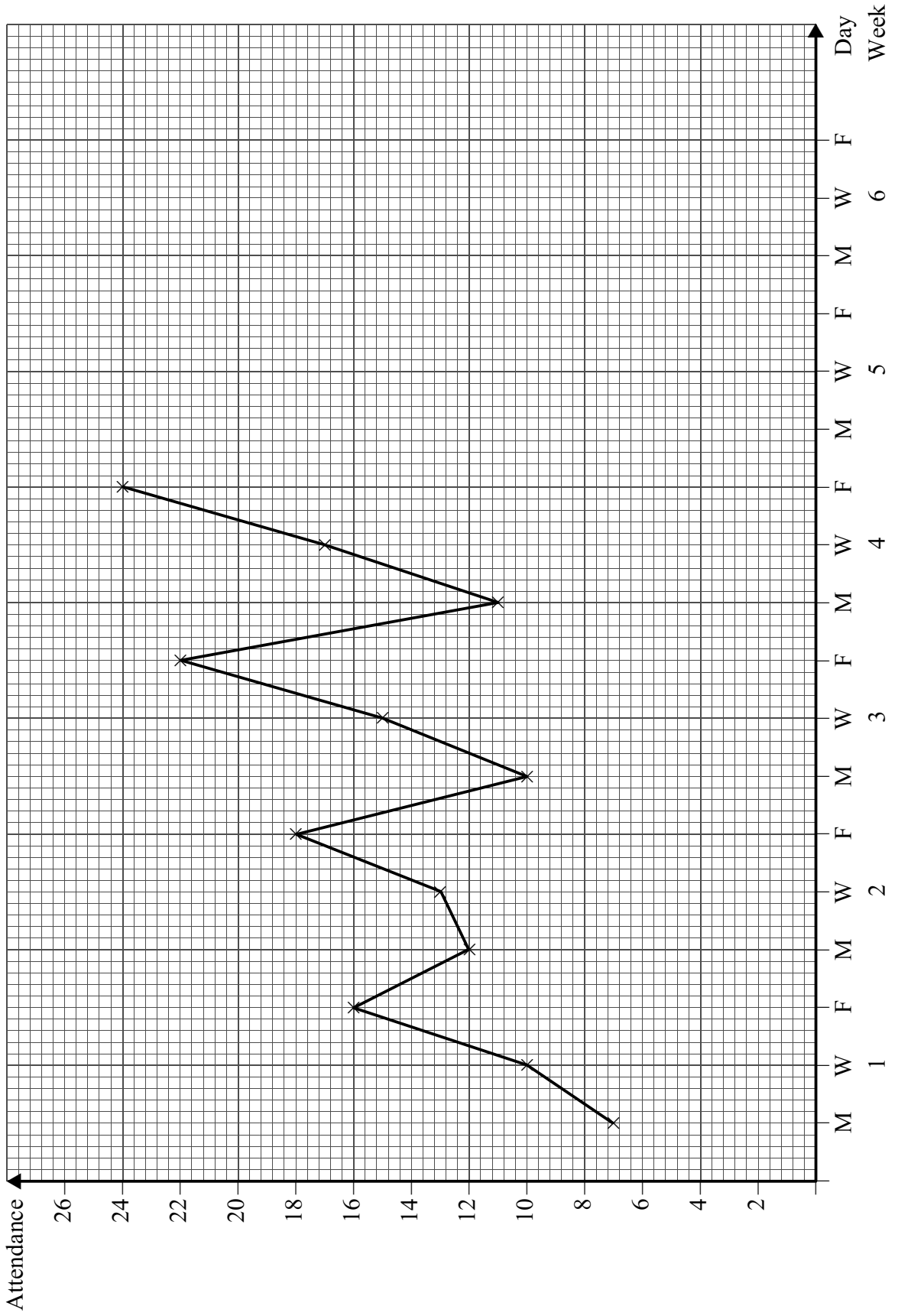**Attendance at Karate Classes**

There is no text printed on this page