AQA

General Certificate of Education
Advanced Subsidiary Examination
June 2015

Statistics   SS02

Unit Statistics 2

Friday 5 June 2015   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 each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
• 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.
• You do not necessarily need to use all the space provided.

Answer **all** questions.

Answer each question in the space provided for that question.

1 Researchers conducted a trial of a new drug intended to reduce pain after surgery. Immediately after their surgery, the patients in the trial received either the usual pain-reduction drug or the new drug. Later, 24 hours after the surgery, the patients recorded their Perceived Pain Index (PPI), where a higher score represents greater pain.

The box plots in **Figure 1** illustrate the data from the trial. The plot for the patients receiving the new drug includes two outliers.

Figure 1
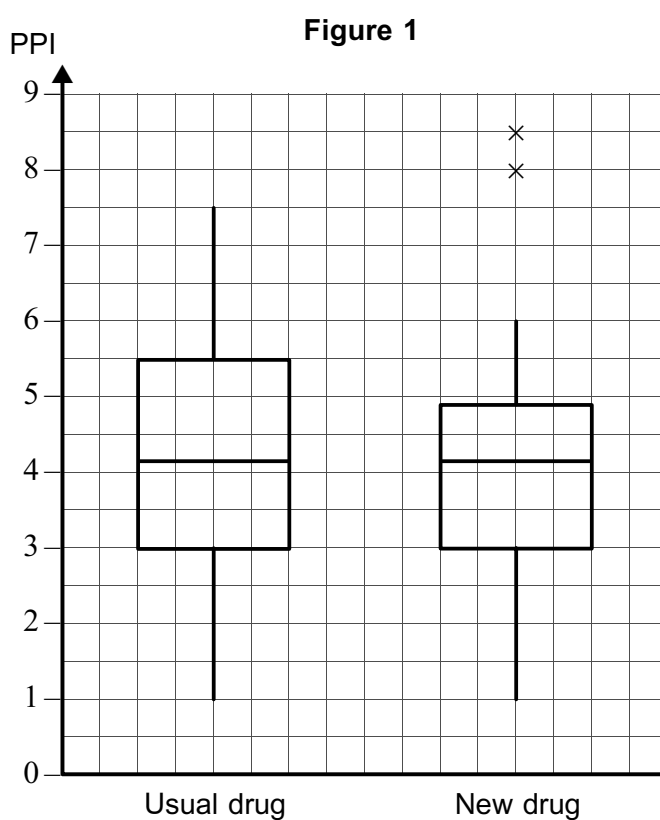
(a) State the highest value of PPI recorded during the trial.

[1 mark]

(b) When comparing the data from the two groups of patients, one of the researchers made the following statement.

"The lowest value of PPI is the same for both drugs, but only the new drug has outliers."

Make **three** further comments comparing the data from the two groups of patients as illustrated in **Figure 1**. Your answer should include at least one similarity and at least one difference.

[3 marks]

QUESTION PART REFERENCE	Answer space for question 1

Turn over ▶

2 **Table 1**, printed below, contains data about UK food and beverage service activities for the years 2008, 2009 and 2010.

Table 1

ANNUAL BUSINESS SURVEY – FOOD & BEVERAGE SERVICE ACTIVITIES

Description	Year	Number of enterprises	Total turnover	Total employment (average)	Total employment costs	Total net capital expenditure
		Number	£ million	Thousand	£ million	£ million
Restaurants and mobile food service activities	2008 2009 2010	63 368 61 192 61 387	22 452 22 658 23 004	753 653 687	6 255 6 460 6 302	1 021 1 104 996
Event catering activities	2008 2009 2010	6 649 6 070 5 767	7 000 6 139 7 008	247 220 254	2 788 2 782 2 957	83 33 117
Other food service activities	2008 2009 2010	887 979 1 264	539 672 682	13 12 18	162 170 192	5 11 18
Beverage serving activities	2008 2009 2010	49 875 45 714 44 351	20 150 19 491 19 574	567 575 543	4 726 4 876 4 542	868 667 375
Total food and beverage service activities	2008 2009 2010	 113 955 112 769	50 141 48 960 50 268	1 578 1 460 1 502	13 930 14 288 13 993	1 978 1 815 1 506

(a) State the total employment costs for people employed in 'Event catering activities' during 2009.

[1 mark]

(b) The figure for the total number of enterprises for 2008 has been omitted. State the value of the missing figure.

[1 mark]

(c) Calculate the **percentage** reduction in total net capital expenditure for 'Beverage serving activities' between 2008 and 2010.

[2 marks]

(d) Calculate the mean employment cost per employee in 'Event catering activities' for 2010.

[2 marks]

QUESTION PART REFERENCE	Answer space for question 2(a), (b), (c), (d)

Question 2 continues on the next page

Turn over ▶

2 (e) (i) On **Figure 2**, printed on the opposite page, draw a line diagram to illustrate the number of enterprises engaged in 'Other food service activities' over the three years.

[1 mark]

(ii) Comment on the trend shown in the line diagram.

[1 mark]

Copy of **Table 1**

ANNUAL BUSINESS SURVEY – FOOD & BEVERAGE SERVICE ACTIVITIES

Description	Year	Number of enterprises	Total turnover	Total employment (average)	Total employment costs	Total net capital expenditure
		Number	**£ million**	**Thousand**	**£ million**	**£ million**
Restaurants and mobile food service activities	2008	63 368	22 452	753	6 255	1 021
	2009	61 192	22 658	653	6 460	1 104
	2010	61 387	23 004	687	6 302	996
Event catering activities	2008	6 649	7 000	247	2 788	83
	2009	6 070	6 139	220	2 782	33
	2010	5 767	7 008	254	2 957	117
Other food service activities	2008	887	539	13	162	5
	2009	979	672	12	170	11
	2010	1 264	682	18	192	18
Beverage serving activities	2008	49 875	20 150	567	4 726	868
	2009	45 714	19 491	575	4 876	667
	2010	44 351	19 574	543	4 542	375
Total food and beverage service activities	2008		50 141	1 578	13 930	1 978
	2009	113 955	48 960	1 460	14 288	1 815
	2010	112 769	50 268	1 502	13 993	1 506

QUESTION PART REFERENCE | **Answer space for question 2(e)**

....

....

....

....

....

....

....

....

....

QUESTION PART REFERENCE	**Answer space for question 2(e)**

Figure 2
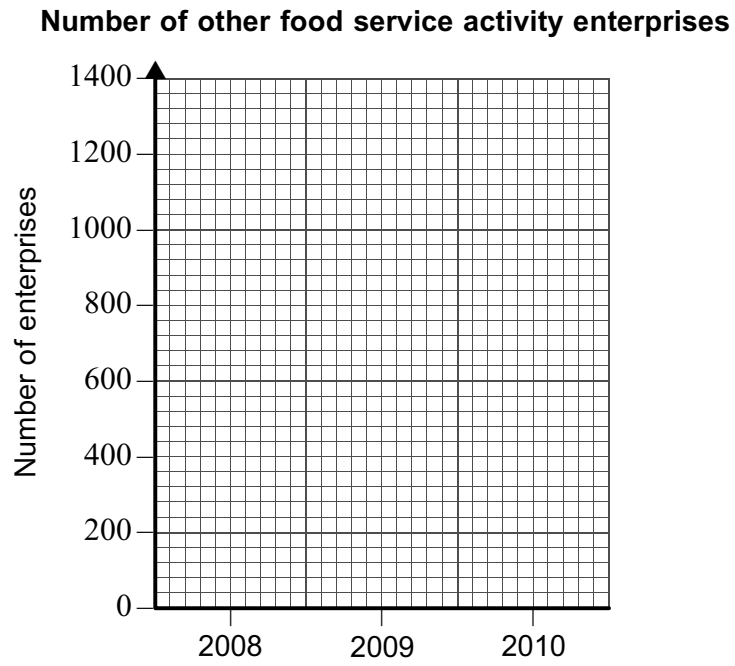
Number of other food service activity enterprises

....

....

....

....

....

....

....

....

....

....

....

....

....

....

Turn over ▶

3 Rika is the manager of a large department store. She believes that the longer their customers stay in the store, the more they are likely to buy. Rika found that customers were spending an average of 24.0 minutes in the store.

In an attempt to encourage customers to stay longer, Rika arranged for relaxing music to be played throughout the store. After the music had been introduced, a sample of 120 customers was observed. The mean time spent in the store by these customers was 25.9 minutes, and the standard deviation was 9.5 minutes.

(a) State an assumption that must be made in order to use this sample as a basis for a hypothesis test.

[1 mark]

(b) Assuming that the assumption that you stated in part **(a)** is valid, test whether the mean time spent in the store by customers has increased. Use the 2% level of significance.

[7 marks]

(c) After the music had been introduced, the mean time spent in the store by customers was, in fact, 25.4 minutes. State, with a reason, whether, in your conclusion in part **(b)**, you made a Type I error, a Type II error or no error.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE	Answer space for question 3

Turn over ▶

4   At a remote hospital, in an area where there are many venomous snakes, the number of patients during one week requiring treatment after a venomous snakebite may be modelled by a Poisson distribution with mean 0.5.

(a)   For this hospital, find the probability that:

(i)   no more than 1 patient requires treatment after a venomous snakebite during a particular week;

[1 mark]

(ii)   at least 5 patients require treatment after a venomous snakebite during a particular period of **8** weeks;

[3 marks]

(iii)   more than 10 patients but fewer than 20 patients require treatment after a venomous snakebite during a particular period of **26** weeks.

[4 marks]

(b)   Each patient who has been bitten by a venomous snake is treated with a single dose of an antivenom which is effective against the venoms of all the snakes common in that area.

The antivenom is expensive and has a limited shelf life, so a delivery of fresh antivenom is made at **4**-week intervals.

The hospital stores just enough antivenom so that the probability that it runs out of antivenom before the next delivery is less than 1 per cent.

Quoting probabilities to justify your answer, state how many doses of antivenom the hospital should have in its store immediately after a delivery of fresh antivenom.

[3 marks]

QUESTION PART REFERENCE	**Answer space for question 4**

QUESTION PART REFERENCE	Answer space for question 4

Turn over ▶

5 The table shows the quarterly numbers of cars first registered in Slovenia from quarter 3 of 2009 to quarter 4 of 2011 together with values of an appropriate n-point moving average.

The data and moving averages are plotted in **Figure 3**, printed on the opposite page.

Year	2009		2010				2011			
Quarter	3	4	1	2	3	4	1	2	3	4
Number of registrations	19 622	18 200	21 112	21 266	18 207	16 409	20 460	20 006	16 764	15 248
Moving average (nearest integer)		20 050	19 696	19 249	19 086	18 771	18 410	18 120		

(a) State the value of n.

[1 mark]

(b) Draw a trend line on **Figure 3**.

[1 mark]

(c) Using your trend line, estimate the seasonal effect for quarter 1.

[3 marks]

(d) Based on these data, forecast the number of registrations for quarter 1 of 2012. Clearly indicate the method you use in order to obtain your forecast.

[3 marks]

(e) The numbers of cars first registered in Slovenia in quarter 1 and in quarter 2 of 2012 were 18 070 and 17 842 respectively.

    (i) Mark these two data points on **Figure 3**.

    (ii) Calculate the next two moving averages and mark these on **Figure 3**.

    (iii) Comment on the trend and on the accuracy of your forecast in part **(d)**.

[6 marks]

QUESTION PART REFERENCE | **Answer space for question 5**

Figure 3
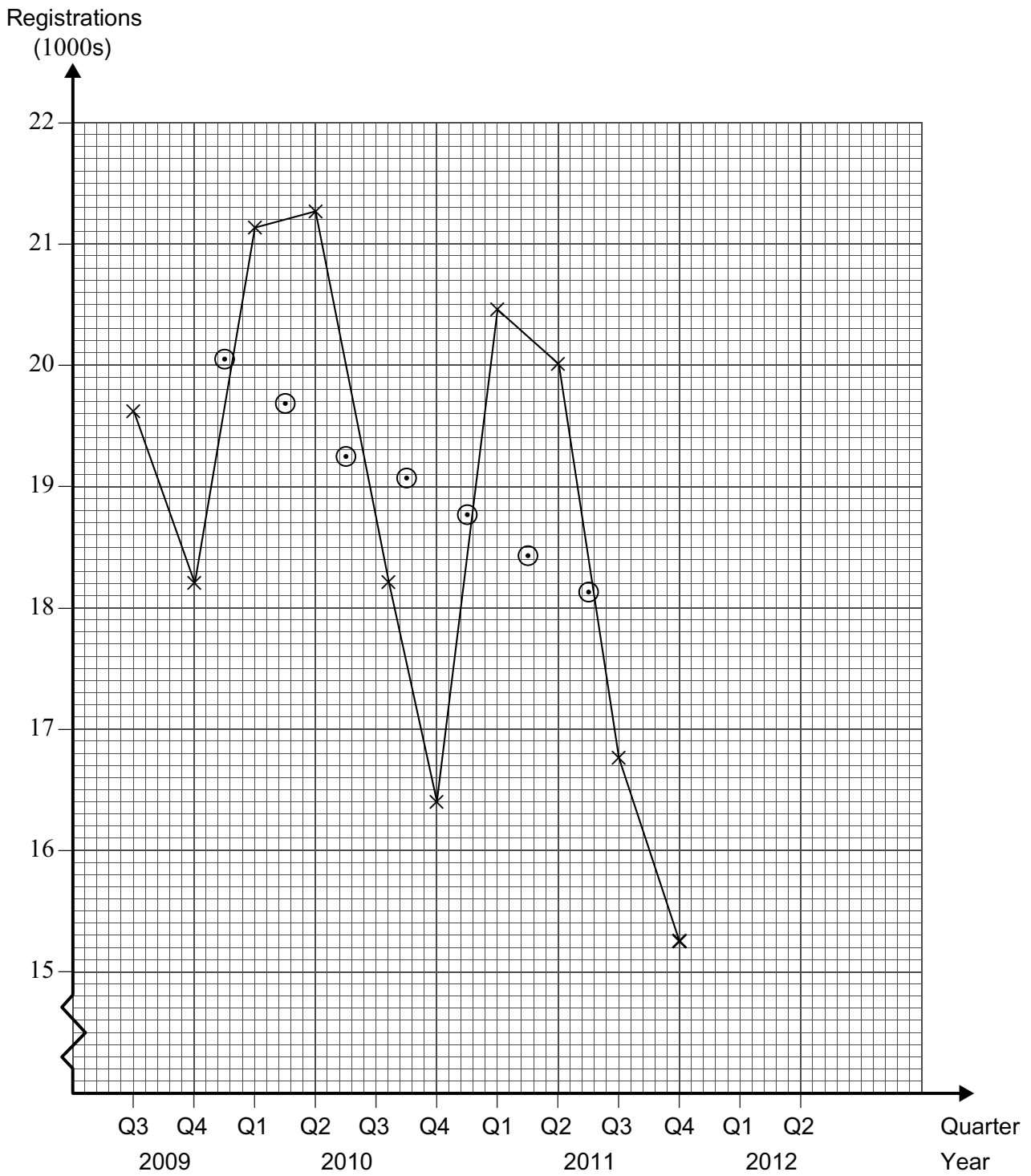
| QUESTION PART REFERENCE | Answer space for question 5 |

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 5

Turn over ▶

6 The Goodwell Medical Practice has a total of 3200 registered patients. The local health authority has asked this practice to complete questionnaires about a sample of 40 patients, but has not said how this sample should be chosen.

The practice has available a list of patients ordered alphabetically by their family name. The patients are numbered in the list from **1** to **3200**. The details of **20** patients are printed on each page of the list.

(a) Dr Dobry suggests picking two pages at random from the list and using the patients on those pages as the sample.

(i) Name this method of sampling.

(ii) Give a reason why Dr Dobry's method would be unlikely to give a representative sample.

[2 marks]

(b) Dr Kalos suggests obtaining the sample of 40 from the list by a systematic method. Explain briefly how this could be done.

[2 marks]

(c) Dr Bueno says the sample should be chosen at random in such a way as to proportionately represent the age distribution of the patients.

(i) Name this method of sampling.

(ii) Given that there are 737 registered patients aged over 60 years, how many of these should be chosen in a sample of 40 chosen using Dr Bueno's method?

[3 marks]

(d) Dr Mabuti suggests obtaining a simple random sample by the following method.

- Obtain a four-digit random number from tables, rejecting any number above 9599.
- Divide the number by 3200 and find the remainder.
- Add 1 to this remainder.
- Select the patient in the list corresponding to this number.
- Carry out this procedure 40 times.

(i) Using this method, which number of the patient in the alphabetical list would be generated by the random number 5817?

(ii) Explain why it is necessary to reject any random number above 9599.

(iii) Explain why it is necessary to add 1.

(iv) Dr Mabuti has omitted one instruction which is needed to make sure this method gives a simple random sample of patients. What instruction needs to be added?

[5 marks]

QUESTION PART REFERENCE	**Answer space for question 6**

Turn over ▶

QUESTION PART REFERENCE	Answer space for question 6

QUESTION PART REFERENCE	Answer space for question 6

Turn over ▶

7 Every Saturday evening, Angus runs a disco at the village hall. The hall must be tidied and cleaned on the morning of the following day, Sunday. This is done by Angus and a variable number of volunteers.

Angus keeps a record of the number of volunteers, X, and the probability distribution for X is given in the table.

x	0	1	2	3	4	5	6 or more
P($X = x$)	p	0.15	0.20	0.21	0.18	0.14	0

(a) (i) Find the value of p.

(ii) Interpret the implication for Angus of this value of p.

[2 marks]

(b) Find the mean value of X and show that, correct to three significant figures, the standard deviation of X is 1.57.

[5 marks]

(c) It is suggested that a Poisson distribution may provide an adequate model for X.

(i) Comment on whether your answers in part **(b)** support this suggestion.

(ii) Give a reason why, **in this context**, a Poisson distribution may not be an appropriate model.

[3 marks]

(d) Every Sunday morning, Angus and any volunteers must also carry 120 chairs into the hall and arrange them for a meeting to be held later that day.

They share this task equally. Hence, the number of chairs, N, which each of them carries into the hall is given by

$$N = \frac{120}{(X + 1)}$$

(i) Find the probability that Angus carries exactly 20 chairs into the hall next Sunday morning.

(ii) Construct a table showing the probability distribution for N.

(iii) Find the mean number of chairs carried into the hall by Angus on a Sunday morning.

[6 marks]

QUESTION PART REFERENCE	**Answer space for question 7**

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 7

Turn over ▶

QUESTION PART REFERENCE	Answer space for question 7

Do not write
outside the
box

QUESTION PART REFERENCE	Answer space for question 7

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED