| 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 2013

# Statistics SS02

## Unit Statistics 2

**Friday 24 May 2013 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.

Do not write outside the box

Answer **all** questions.

Answer each question in the space provided for that question.

**1** Chris works as the receptionist at the Zubar Dental Practice. The number of requests from male patients for emergency dental treatment that she receives during any one day may be modelled by a Poisson distribution with mean 2.6.

**(a)** Find the probability that the number of requests from male patients for emergency dental treatment that Chris receives on a particular day is:

**(i)** exactly 4; *(2 marks)*

**(ii)** at least 1. *(2 marks)*

**(b)** Find the probability that the number of requests from male patients for emergency dental treatment that Chris receives during a period of **5 days** is 15 or fewer. *(2 marks)*

**(c)** The number of requests from **female** patients for emergency dental treatment that Chris receives during any one day is independent of the number of such requests from male patients and may be modelled by a Poisson distribution with mean 3.4. The Zubar Dental Practice can provide emergency treatment to a **total** of 8 patients each day.

Find the probability that Chris receives more requests for emergency dental treatment in one day than the Zubar Dental Practice can provide on that day. *(3 marks)*

QUESTION PART REFERENCE

**Answer space for question 1**

*Do not write outside the box*

*QUESTION PART REFERENCE*

**Answer space for question 1**

Do not write outside the box

**2** A market trader sells bags of cherries. A sign on his stall says, '1 lb bags of cherries'. The unit of weight 1 lb is equal to 453.6 grams. Sophie, the market inspector, suspects that the bags may, on average, contain less than 453.6 grams.

Sophie asks her assistant, Kevin, to investigate her suspicion. She tells Kevin that, from previous measurements, the weights of bags of cherries from this trader may be assumed to have a normal distribution with standard deviation 10 grams. Kevin weighs 6 bags of cherries and obtains the following weights in grams.

448.2 461.9 455.8 437.0 442.5 441.4

**(a)** Assuming that the distribution of weights of bags of cherries is still normal with standard deviation 10 grams, investigate Sophie's suspicion at the 10% significance level. *(7 marks)*

**(b)** Sophie discovers that Kevin has included the weight of the bag in his data, as well as the weight of the cherries. Each bag weighs 5 grams. With this information, show that, on the basis of the above data, Sophie can now confirm her suspicion at the **1%** significance level. *(4 marks)*

**(c)** Kevin suggests that, in rejecting the trader's claim, Sophie may have made a Type I error or a Type II error.

State, giving a reason, which type of error (Type I or Type II) Sophie may have made in this case. *(2 marks)*

QUESTION PART REFERENCE

**Answer space for question 2**

Do not write outside the box

QUESTION PART REFERENCE

**Answer space for question 2**

Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 2

Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 2

Do not write outside the box

**3** A bank has an ATM (Automated Teller Machine) which customers can use to withdraw cash. Withdrawals can be made in fixed amounts, £$X$. The table shows the amounts available and the probability distribution for $X$.

| $x$ | P($X = x$) |
|---|---|
| 10 | 0.18 |
| 20 | 0.44 |
| 50 | 0.13 |
| 100 | 0.08 |
| 200 | 0.17 |

**(a) (i)** Find the value of $\mathrm{E}(X)$. *(2 marks)*

**(ii)** Show that the standard deviation of $X$ is 68.0, correct to three significant figures. *(3 marks)*

**(iii)** Find the probability that at least one out of a random sample of three customers withdraws more than the mean amount of cash. *(3 marks)*

**(b)** The bank is considering making an additional amount of £300 available. It is expected that some of the customers who currently withdraw £200 would then withdraw £300.

State whether this change would increase, decrease or leave unchanged:

**(i)** the mean of $X$;

**(ii)** the standard deviation of $X$. *(2 marks)*

**(c)** A small number of customers use the ATM to see the balance in their account and do not withdraw any cash. If the table were changed to include these customers, **explain why** this would decrease the mean of $X$. *(2 marks)*

QUESTION PART REFERENCE

**Answer space for question 3**

Do not write outside the box

QUESTION PART REFERENCE

**Answer space for question 3**



Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 3

Do not write outside the box

QUESTION PART REFERENCE

**Answer space for question 3**

Do not write outside the box

**4** The numbers of customers at Carlo's restaurant on the Fridays, Saturdays and Sundays during November 2012 are shown in the table, together with an $n$-point moving average.

| **Date** | **2** | **3** | **4** | **9** | **10** | **11** | **16** | **17** | **18** | **23** | **24** | **25** | **30** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Day** | **Fr** | **Sa** | **Su** | **Fr** | **Sa** | **Su** | **Fr** | **Sa** | **Su** | **Fr** | **Sa** | **Su** | **Fr** |
| **Number** | 71 | 77 | 59 | 77 | 86 | 66 | 76 | 95 | 85 | 94 | 104 | 82 | 102 |
| **Moving average** | | 69 | 71 | 74 | 76.3 | 76 | 79 | 85.3 | 91.3 | 94.3 | 93.3 | $a$ | |

These data have been plotted on the graph opposite, along with a trend line.

**(a) (i)** State the value of $n$. *(1 mark)*

**(ii)** Calculate the value of the missing moving average, $a$, and plot this value on the graph. *(2 marks)*

**(b) (i)** Estimate the seasonal effect for Saturday. *(3 marks)*

**(ii)** Hence forecast the number of customers on Saturday 1 December 2012. *(3 marks)*

**(iii)** State one reason why the data for November may not provide an accurate forecast for the number of customers at Carlo's restaurant on each Friday, Saturday and Sunday during December. *(1 mark)*

**(c)** During November, there was bad weather on one of the days listed in the table, making travel difficult. Also, on a different day listed in the table, Carlo ran a promotion that offered, 'Free glass of wine or soft drink with each meal'.

Using the graph, state, with a reason, on which date it is likely that:

**(i)** the bad weather occurred; *(2 marks)*

**(ii)** the promotion was offered. *(2 marks)*

QUESTION PART REFERENCE

**Answer space for question 4**

.....

.....

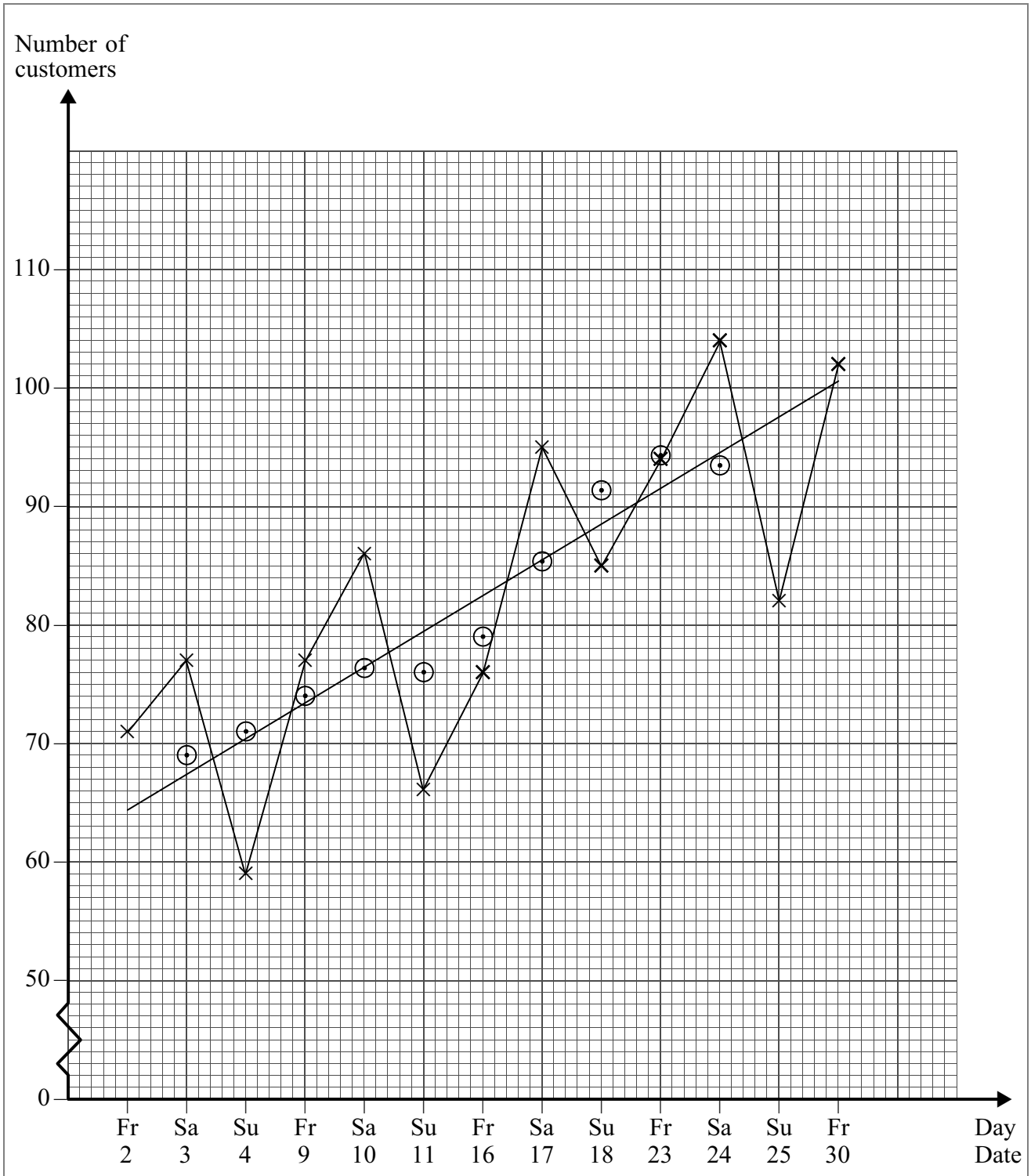
.....

.....

.....

.....

Do not write outside the box




QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE

**Answer space for question 4**

Do not write outside the box

**5** A field study centre is near to a stream of length 1.8 km. The tutors at the centre divide this length into 20 m sections, providing 90 sections where students visiting the centre can collect data about the flow of water in the stream. The sections are numbered from 1 to 90. For the first 840 m, the stream is classed as a 'first order' stream; for the next 360 m, it is classed as a 'second order' stream; and for the remaining sections, it is classed as a 'third order' stream.

A school party visiting the centre is divided into 5 teams of students who are to investigate how the flow of water varies along the stream. Each team will collect data at 3 sections, so 15 different sections must be selected. The collection of data by the teams will be supervised by two teachers from the school and one tutor from the centre.

The students are asked to suggest how the sample of 15 sections should be selected.

**(a)** Anders suggests selecting a simple random sample of 15 sections, using random numbers from tables.

**(i)** Describe, in detail, how this might be done.

**(ii)** State two possible disadvantages, one statistical and one practical, of collecting data from sections selected in this way. *(6 marks)*

**(b)** Barbara suggests rolling a dice to choose the number of the first section to be used, and then selecting every sixth section after that to complete the sample.

**(i)** Name this type of sampling.

**(ii)** State, with a reason, whether a sample obtained in this way will be **random**.

**(iii)** State, with a reason, whether a sample obtained in this way will be **stratified**. *(6 marks)*

**(c)** Caleb suggests choosing a block of 5 consecutive sections along the first order stream and giving one section to each team, then repeating this for the second order and for the third order streams.

**(i)** Name this type of sampling.

**(ii)** State **two** possible advantages of collecting data from sections selected in this way.

**(iii)** State **one** possible disadvantage of collecting data from sections selected in this way. *(4 marks)*

Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

**Answer space for question 5**

Do not write outside the box

QUESTION PART REFERENCE

**Answer space for question 5**

Do not write outside the box

**6** **Table 1** shows the areas in Canada planted with field crops and special crops during the period 2007 to 2011.

**Table 1**

| **Field and special crops** (Seeded area in thousands of hectares) | | | | | |
|---|---|---|---|---|---|
| **Year** | **2007** | **2008** | **2009** | **2010** | **2011** |
| **Field crops** | | | | | |
| All wheat | 8 849.5 | 10 192.4 | 10 065.3 | 8 549.6 | 8 718.1 |
| Canola | 6 376.2 | 6 539.6 | 6 687.3 | 7 125.8 | 7 633.2 |
| Barley | 4 396.8 | 3 786.6 | 3 505.9 | 2 796.6 | 2 619.1 |
| Oats | 2 188.4 | 1 758.4 | 1 510.1 | 1 219.3 | 1 258.0 |
| Flaxseed | 528.0 | 631.3 | 692.0 | 374.3 | 281.2 |
| Rye | 167.9 | 168.0 | 167.9 | 131.5 | 105.3 |
| Soybeans | 1 180.1 | 1 202.4 | 1 395.3 | 1 483.0 | 1 549.9 |
| Corn for grain | 1 391.5 | 1 204.0 | 1 203.5 | 1 214.3 | 1 217.7 |
| Tame hay | 8 239.2 | 8 201.6 | 8 183.1 | 8 168.3 | 7 967.4 |
| **Special crops** | | | | | |
| Canary seed | 178.1 | 167.9 | 149.8 | 159.8 | 95.1 |
| Lentils | 580.8 | 706.2 | 971.3 | 1 408.3 | 1 040.0 |
| Sunflower seed | 80.9 | 68.8 | 64.7 | 54.6 | 14.2 |
| Mustard seed | 186.2 | 194.2 | 212.4 | 194.2 | 127.5 |
| Dry peas | 1 469.0 | 1 616.6 | 1 521.7 | 1 466.9 | 942.0 |

Source: *Statistics Canada*, 2011

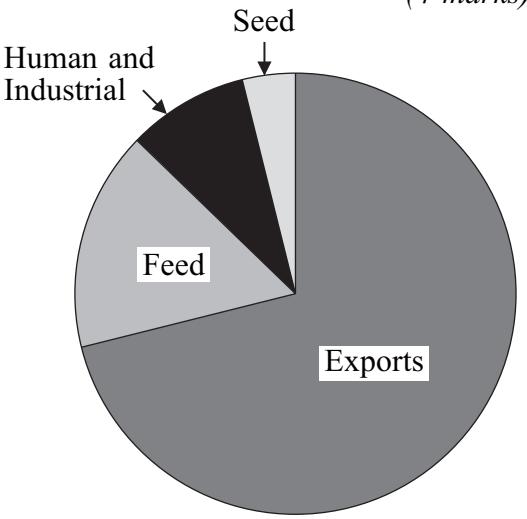
**(a)** Find the difference between the seeded area of dry peas and the seeded area of lentils in Canada in 2008. *(2 marks)*

**(b)** Find the mean seeded area of flaxseed in Canada during the period 2007 to 2011 inclusive. *(2 marks)*

**(c)** You may assume that fields planted with wheat produce an average of 2.2 tonnes of wheat per hectare. The pie chart and **Table 2** below indicate the uses to which Canadian wheat is put. Using the information provided in **Table 1**, **Table 2** and the pie chart, estimate, to the nearest 100 000 tonnes, the number of tonnes of Canadian wheat exported during 2011. *(4 marks)*

**Table 2**

| Use | Angle |
|---|---|
| Exports | 256° |
| Feed | 58° |
| Human and Industrial | 32° |
| Seed | 14° |



Do not write outside the box

**(d)** Three scatter diagrams were drawn of the data for Canada for the years 2007 to 2011, in each case with seeded area of barley on the horizontal axis and a different field crop on the vertical axis. **Figures 1**, **2** and **3** show sketches of the regression lines obtained. The three field crops were oats, soybeans and corn for grain.

**Figure 1**

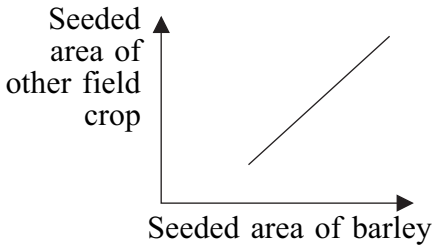



**Figure 2**

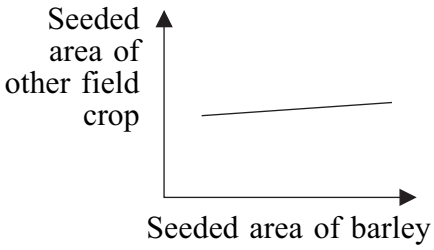
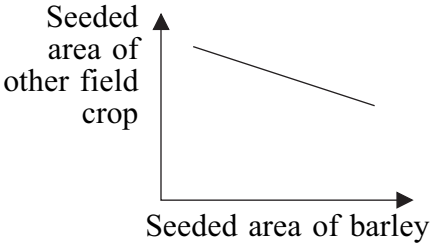



**Figure 3**




By considering the data in **Table 1**, state, with a reason, which of the three field crops, oats, soybeans or corn for grain, is most likely to have been on the vertical axis in:

**(i)** **Figure 1**;

**(ii)** **Figure 2**;

**(iii)** **Figure 3**. *(3 marks)*

QUESTION PART REFERENCE

**Answer space for question 6**

Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 6

Do not write outside the box

QUESTION PART REFERENCE

## Answer space for question 6

**END OF QUESTIONS**

*Do not write outside the box*

**There are no questions printed on this page**

**DO NOT WRITE ON THIS PAGE**
**ANSWER IN THE SPACES PROVIDED**

Copyright © 2013 AQA and its licensors. All rights reserved.