Paper Reference(s)

1389/1H

Edexcel GCSE

Statistics

Paper 1H

Higher Tier

Tuesday 16 June 2009 – Morning

Time: 2 hours 30 minutes

Materials required for examination
Ruler graduated in centimetres and
millimetres, protractor, compasses,
pen, HB pencil, eraser,
electronic calculator.

Items included with question papers
Nil

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions. Write your answers in the spaces provided in this question paper.
You must NOT write on the formulae page or any blank pages. Anything you write on these pages will gain NO credit.
If you need more space to complete your answer to any question, use additional answer sheets.

Information for Candidates

The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
This question paper has 8 questions in Section A and 7 questions in Section B. The total mark for this paper is 100.
There are 28 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

Show all stages in any calculations.
Work steadily through the paper. Do not spend too long on one question.
If you cannot answer a question, leave it and attempt the next one.
Return at the end to those you have left out.

Turn over

GCSE Statistics 1389

Higher Tier Formulae

You must not write on this page.
Anything you write on this page will gain NO credit.

Mean of a frequency distribution $\qquad = \dfrac{\sum fx}{\sum f}$

Mean of a grouped frequency distribution $\qquad = \dfrac{\sum fx}{\sum f}$, where x is the mid-interval value.

Variance $\qquad = \dfrac{\sum (x - \bar{x})^2}{n}$

Standard deviation (set of numbers) $\qquad \sqrt{\left[\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n} \right)^2 \right]}$

or $\qquad \sqrt{\left[\dfrac{\sum (x - \bar{x})^2}{n} \right]}$

where $\bar{x}$ is the mean set of values.

Standard deviation (discrete frequency distribution) $\qquad \sqrt{\left[\dfrac{\sum fx^2}{\sum f} - \left(\dfrac{\sum fx}{\sum f} \right)^2 \right]}$

or $\qquad \sqrt{\left[\dfrac{\sum f(x - \bar{x})^2}{\sum f} \right]}$

Spearman's Rank Correlation Coefficient $\qquad 1 - \dfrac{6\sum d^2}{n(n^2 - 1)}$

SECTION A

Answer ALL the questions. Write your answers in the spaces provided.

You must write down all stages in your working.

1. The pie chart shows information about the numbers of officers in the Royal Navy, Army and Air Force.

    Officers in the Royal Navy, Army and Air Force

    Royal Navy

    Army

    Air Force

    Total: 30 000 officers
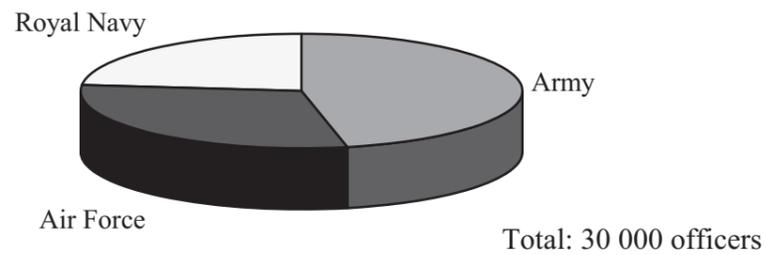
    The pie chart is misleading.

    (a) Write down one reason why.

    ....

    ....

    ....

    (1)

    A new pie chart is to be drawn which is not misleading.

    The number of officers in the Royal Navy is 7000

    (b) Calculate the angle needed for the Royal Navy.

    ....°

    (2)

2. The bar chart shows the number of repeat programmes shown by BBC1, BBC2 and ITV1 during one week.

Repeat programmes

(Source: Radio Times)
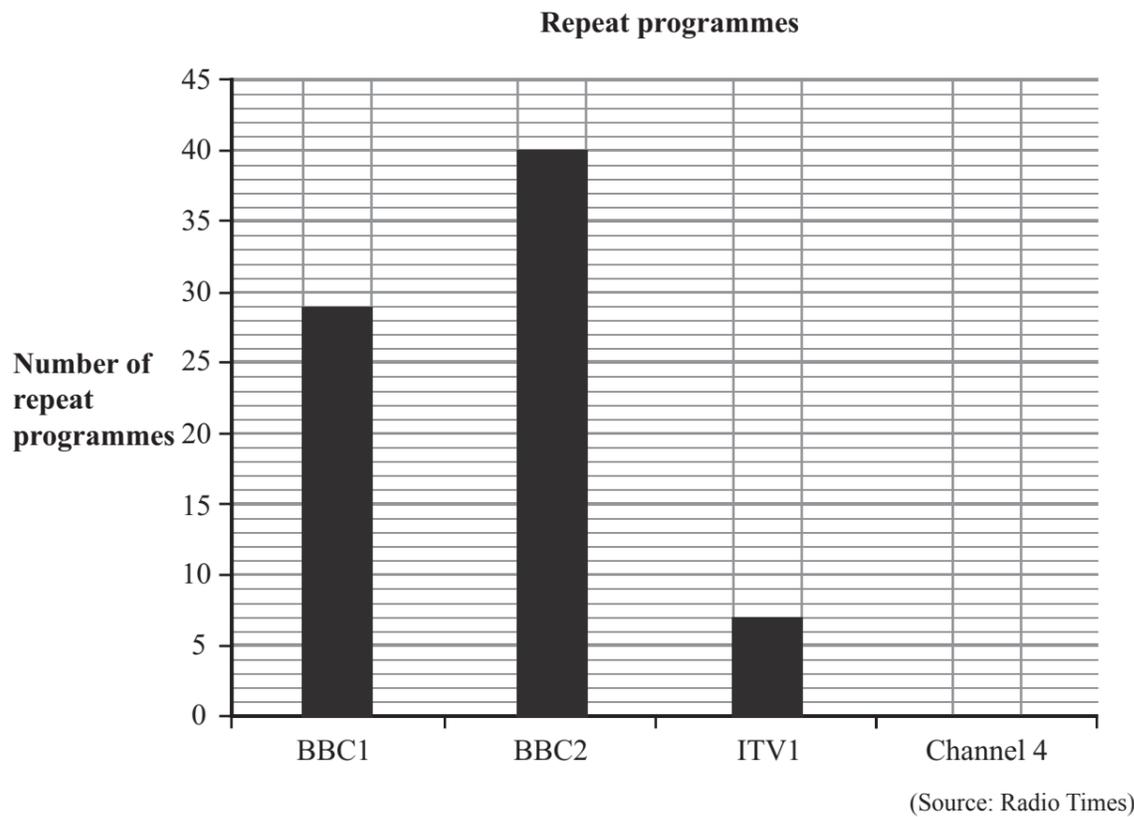
During the same week Channel 4 showed 39 repeat programmes.

(a) Complete the bar chart to show this information.

(1)

(b) Work out the **total** number of repeat programmes shown during this week by BBC1, BBC2, ITV1 and Channel 4

....

(2)

(c) Compare the numbers of repeat programmes on BBC1, BBC2, ITV1 and Channel 4

....

....

....

....

(2)

(Total 5 marks)

3. Here is a random number table.

```
5  1      1  9      5  3      8  4      3  8
6  3      0  9      9  3      1  1      5  4
3  1      9  3      9  0      6  6      6  5
5  6      9  2      1  5      4  6      9  3
9  2      1  1      2  7      0  6      0  9
```

Michael wants to take a random sample of **six** people from a list of 60 people.

He starts from the top left of the table then works across the table row by row.

(a) Write down the set of numbers Michael could use.

....

....

(2)

(b) Explain how Michael could use these numbers to obtain his sample.

....

....

....

....

....

(2)

(Total 4 marks)

Q3

4. Life expectancy is the number of further years that a person is expected to live.

Example: A male was aged 20 in 1997
He is expected to live another 55.4 years.

The table gives information about the life expectancy for people in the United Kingdom in the years 1997 to 2004

Life Expectancy in the United Kingdom										
	Males						Females			
	Age						Age			
Year	At Birth	20	30	60	80		At Birth	20	30	60	80
1997	74.5	55.4	45.9	18.8	6.7		79.6	60.2	50.4	22.5	8.4
1998	74.8	55.6	46.1	18.9	6.7		79.7	60.4	50.5	22.6	8.4
1999	75.0	55.9	46.3	19.2	6.8		79.9	60.5	50.7	22.8	8.5
2000	75.4	56.2	46.6	19.5	7.0		80.2	60.8	51.0	23.0	8.6
2001	75.7	56.5	46.9	19.8	7.1		80.4	61.0	51.2	23.2	8.7
2002	75.9	56.7	47.2	20.0	7.2		80.5	61.1	51.3	23.3	8.7
2003	76.3	57.0	47.4	20.2	7.3		80.7	61.3	51.5	23.4	8.7
2004	76.6	57.4	47.8	20.5	7.4		81.0	61.5	51.7	23.6	8.8

(Source: Statistics.gov.uk)

A female was aged 30 in 2002

(a) Write down her life expectancy.

.... years

(1)

(b) Discuss how the life expectancy of a male compares with the life expectancy of a female of the same age.

....

....

....

....

....

(2)

(c) (i) Describe how the life expectancy for both males and females has changed between the years 1997 and 2004

....

....

....

....

(ii) Give **one** possible reason for this.

....

....

....

....

(2)　　Q4

(Total 5 marks)

5. The time series graph shows information about the number of sight tests in England in each year from 1995 to 2004

Sight tests in England
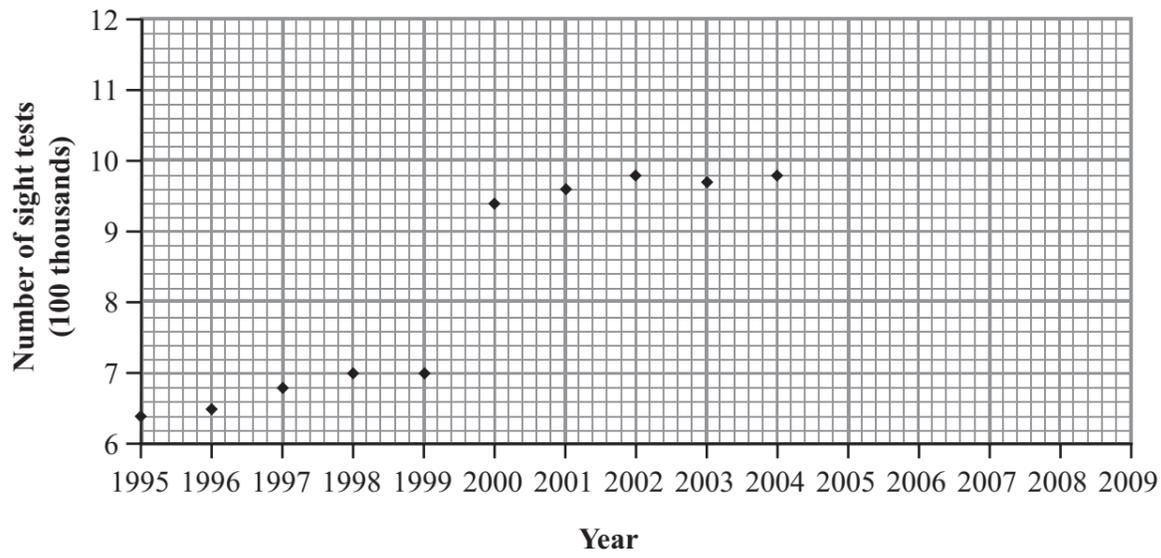
(Source: www.dh.gov.uk)

(a) Explain what the time series graph shows about the number of sight tests from 1995 to 2004

....

....

....

(2)

(b) Use the information in the time series graph from 2000 to 2004 to predict the number of sight tests in 2009

....

(1)

(c) Comment on the reliability of your prediction.

....

....

....

....

(2)

(Total 5 marks)

6. Helen and Jan are playing badminton. Helen's odds for winning a game are 3:2 on.

(a) Write down the odds against Helen winning the next game.

....
(1)

(b) Write down the probability that Jan will win the next game.

....
(1)

Helen and Jan play 20 games of badminton.

(c) Work out the number of games that Jan can expect to win.

....
(1)

(Total 3 marks)

7. Talil took a History exam.
The marks in this exam are normally distributed with a mean mark of 65 and a standard deviation of 10

Talil had a mark of 52 in the History exam.

(a) Calculate Talil's standardised score for the History exam.

....

(3)

Talil also took a Science exam.
His standardised score for the Science exam was -0.25

(b) Did Talil do better in the History exam or in the Science exam?
Give a reason for your answer.

....

....

....

(2)

(Total 5 marks)

8. Stephanie recorded the time she took to travel to work on each of 50 days.
The table shows information about these times.

Time (x minutes)	Frequency (f)
$20 < x \leqslant 30$	4
$30 < x \leqslant 38$	9
$38 < x \leqslant 42$	12
$42 < x \leqslant 50$	18
$50 < x \leqslant 60$	7

(a) Calculate an estimate of the mean time Stephanie took to travel to work.

.... minutes

(3)

(b) Calculate an estimate of the standard deviation of these times.

You may use $\sum fx^2 = 91\,367$

.... minutes

(2) **Q8**

(Total 5 marks)

TOTAL FOR SECTION A: 35 MARKS

H 3 4 3 6 1 A 0 1 1 2 8

SECTION B

Answer ALL the questions. Write your answers in the spaces provided.

You must write down all stages in your working.

1. The manager of a factory wants to carry out a survey to find out the workers' views on selling them overalls.

    There are 1200 workers in the factory.

    (a) Give **two** reasons why the manager would take a sample rather than a census.

    1 ....

    ....

    ....

    2 ....

    ....

    ....

    (2)

    The manager wants the sample to be made up of both men and women.

    There are 700 men and 500 women.

    (b) Write down the name of a sampling process that would be suitable.

    ....

    (1)

    The manager has decided to use a questionnaire to find out the workers' views on selling them overalls.

    (c) Three of the questions are shown below.
         Write down **one** criticism of each question.

    Question 1 'Give one reason why you support the idea of having overalls sold at the factory.'

    Criticism ....

    ....

    ....

    ....

    ....

Question 2 'How much would you be willing to pay for your overalls?' Please tick one box.

Less than £10   £10 to £20   £15 to £30   Over £30

Criticism ....

....

....

....

....

Question 3 'What colour should the overalls be?'

Criticism ....

....

....

....

....

(3)

The manager decided to do a pilot study.

(d) Write down **two** reasons for doing a pilot study.

1 ....

....

....

2 ....

....

....

(2) | **Q1**

(Total 8 marks)

**2.** A scientist recorded the temperature of boiling water at 7 different heights above sea level.

The table shows the temperature of boiling water, y °C, at a height, x metres, above sea level.

Height above sea level (x metres)	1000	1600	2200	3600	4800	6400	7200
Temperature of boiling water (y °C)	96	94	92	88	84	78	76

(Source: www.wonderquest.com)

(a) Draw a scatter diagram for the information given in the table.

Temperature of boiling water at different heights above sea level
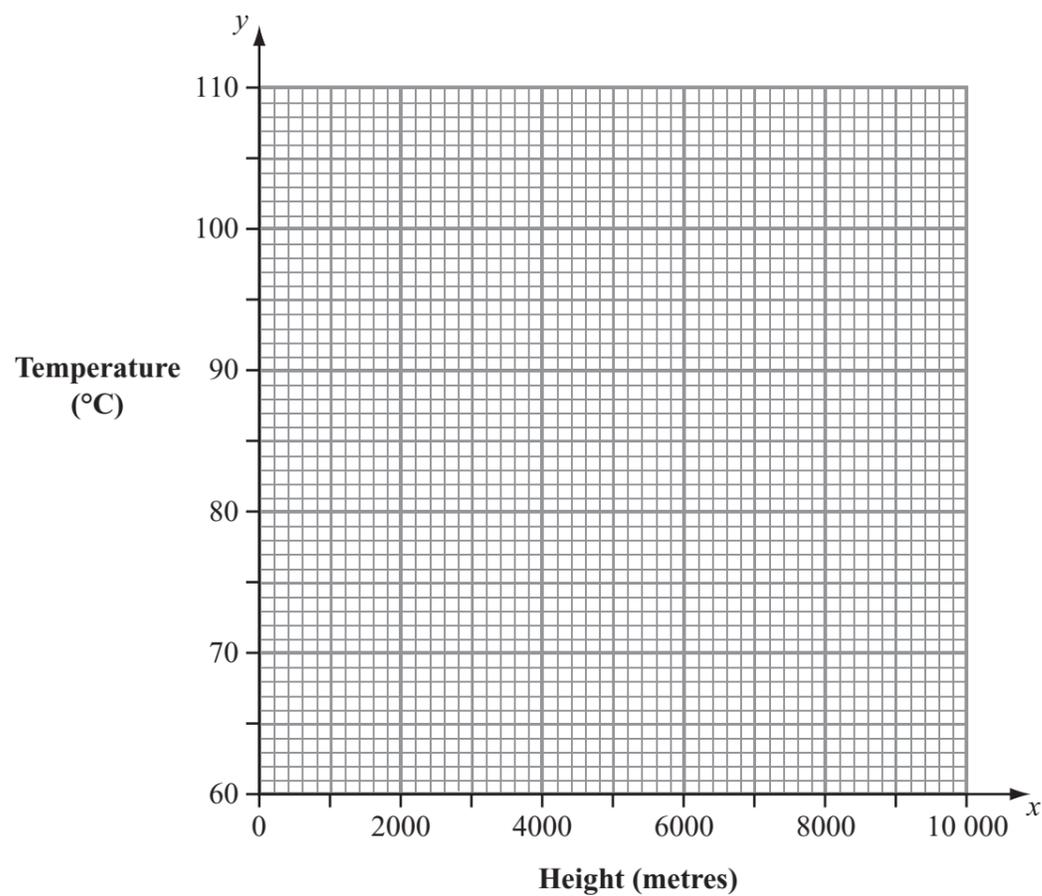
(2)

(b) Describe and interpret the correlation between the temperature of boiling water and the height above sea level.

....

....

....

(2)

(c) On the scatter diagram, draw a line of best fit.

(1)

On Mount Everest there is a base camp at a height of 5600 metres above sea level.

(d) Use your line of best fit to estimate the temperature of boiling water at this base camp.

.... °C

(1)

(e) How reliable is this estimate?
Give a reason for your answer.

....

....

....

(1)

An equation of a line of best fit for these data is $y = -\dfrac{1}{300}x + 100$

(f) Give a practical interpretation of the meaning of $-\dfrac{1}{300}$

....

....

(2) Q2

(Total 9 marks)

3. The table shows the number of police and the number of prisoners in each of seven countries.

Country	Number of police	Number of prisoners				
Sri Lanka	34 609	19 974				
Romania	48 842	41 817				
Canada	56 020	36 024				
Kazakhstan	69 096	58 300				
Colombia	91 155	54 034				
Poland	100 778	80 093				
Spain	155 656	59 251				

(Source: www.nationmaster.com)

The data in the table was collected from the Internet.

(a) Write down **one** advantage and **one** disadvantage of collecting data from the Internet.

Advantage ....

....

Disadvantage ....

....

(2)

(b) Work out Spearman's rank correlation coefficient for these data.

....

(3)

(c) Interpret your answer to part (b).

....

....

....

(2)

(Total 7 marks)

4. The table gives information about the numbers of registered vehicles in Great Britain in 1950 and in 2000

	Cars	Motor cycles	Other vehicles	Total (thousands)
1950	49.8%	34.0%	16.2%	3970
2000	80.2%	16.9%	2.9%	28 898

(Source: www.dft.gov.uk)

(a) Use information from the table to complete the composite bar chart.

Percentages of registered vehicles in Great Britain

Key	
Cars	
Motor cycles	
Other vehicles	
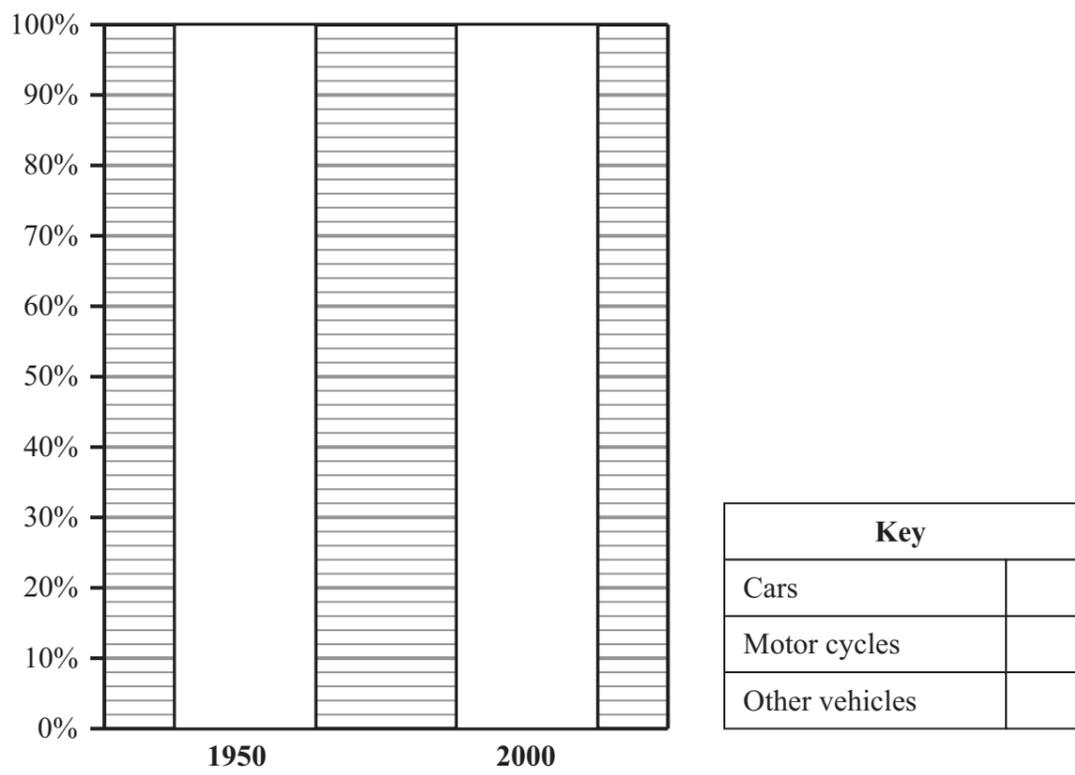
(3)

Bill says that there were more motorcycles in 1950 than in 2000

(b) Write down whether or not Bill is right.
Give a reason for your answer.

....

....

....

(2)

The table shows information about the number of registered vehicles for each of the years 2001 to 2005

	2001	2002	2003	2004	2005
Number of registered vehicles (thousands)	29 747	30 557	31 207	32 259	32 897

(c) Calculate the chain base index numbers for 2002, 2003, 2004 and 2005
     Give your answers correct to 1 decimal place.

2002 = ...

2003 = ...

2004 = ...

2005 = ...

(3)

(d) Calculate the geometric mean for your 4 chain base index numbers.

....

(2)

(e) Interpret your answer to part (d).

....

....

(2)   Q4

(Total 12 marks)

H34361A01928

5. Karl collected information about the Paralympic World Cup 2007

He collected the times for 31 male swimmers in the 50 m Freestyle event.

The stem and leaf diagram shows this information.

The times are given to the nearest second.

Male swimmers' times

```
2 | 8 8 8 8 9 9
3 | 0 2 2 2 2 2 3 3 3 3 4 4 4 6 6 6 9 9
4 | 0 1 1 2 7
5 | 4 8
```
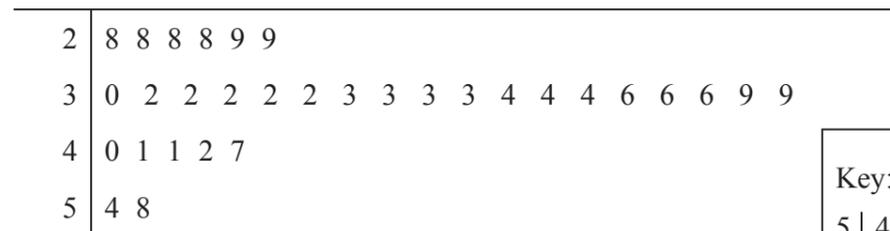
Key:
5 | 4 = 54 seconds

(Source: www.visaparalympicworldcup.com)

(a) Use the information in the stem and leaf diagram to complete the following table.

Lowest value	
Lower quartile	32
Median	
Upper quartile	39
Highest value	

(3)

(b) Identify any outliers for the times of the male swimmers.

....

(3)

He collected the times for female swimmers in the 50 m Freestyle event.

The box plot shows information about these times.

50 m Freestyle event

Female swimmers

Male swimmers

20   30   40   50   60   70

Time (seconds)
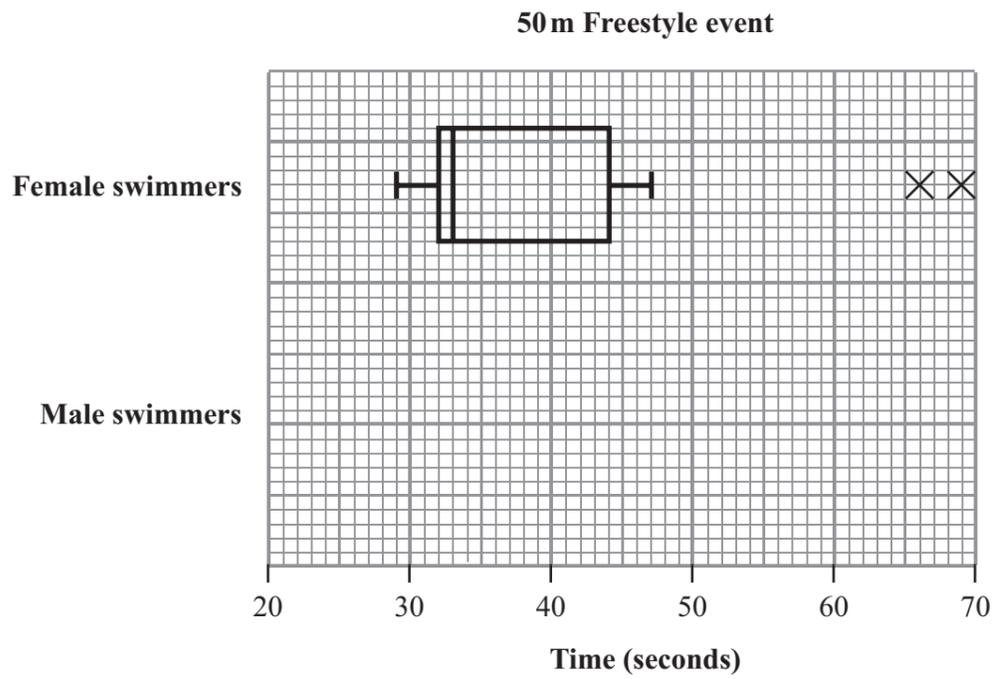
(c) On the grid, draw a box plot to show the distribution of the times of the male swimmers.

(3)

(d) Compare the distributions of the times of the female swimmers and the times of the male swimmers.

....

....

....

....

....

....

(3)

Q5

(Total 12 marks)

6. A machine fills packets with flour.

The label on each packet says the contents weigh 1500g.

(a) Give **one** reason why it may not be practical to check the weight of each packet.

....

....

(1)

Here is a list of words.

continuous   discrete   nominal   categorical   numerical   rank

(b) Use **two** words from the list to complete the following sentence.

The weights of packets of flour are .... and .... data.

(2)

At regular intervals a sample of 10 packets is taken from the machine.

The sample mean is calculated.

The sample means are normally distributed with a mean of 1520 g and a standard deviation of 4 g.

(c) Between what limits would you expect 95% of the sample means to lie?

....

(2)

Martha takes a sample of 10 packets. She finds the weight, in grams, of each packet.

Here are her results.

1518   1526   1535   1528   1532   1522   1533   1536   1527   1533

(d) Calculate the mean of this sample.

.... g

(1)

The allowable limits for the sample means are $1520\,\text{g} \pm 3$ standard deviations.

(e) Suggest what action Martha should take.
   Give a reason for your answer.

....

....

....

(3)

(Total 9 marks)

Q6

7. The probability that Abena has a sickle cell gene is $\frac{1}{10}$

The probability that Kofi has a sickle cell gene is $\frac{1}{5}$

These probabilities are independent.

(Source: www.sicklecell.md)

(a) Complete the probability tree diagram.

Abena	Kofi

$\frac{1}{5}$

Has gene

$\frac{1}{10}$ Has gene

.... Does not have gene

.... Has gene

.... Does not have gene

.... Does not have gene
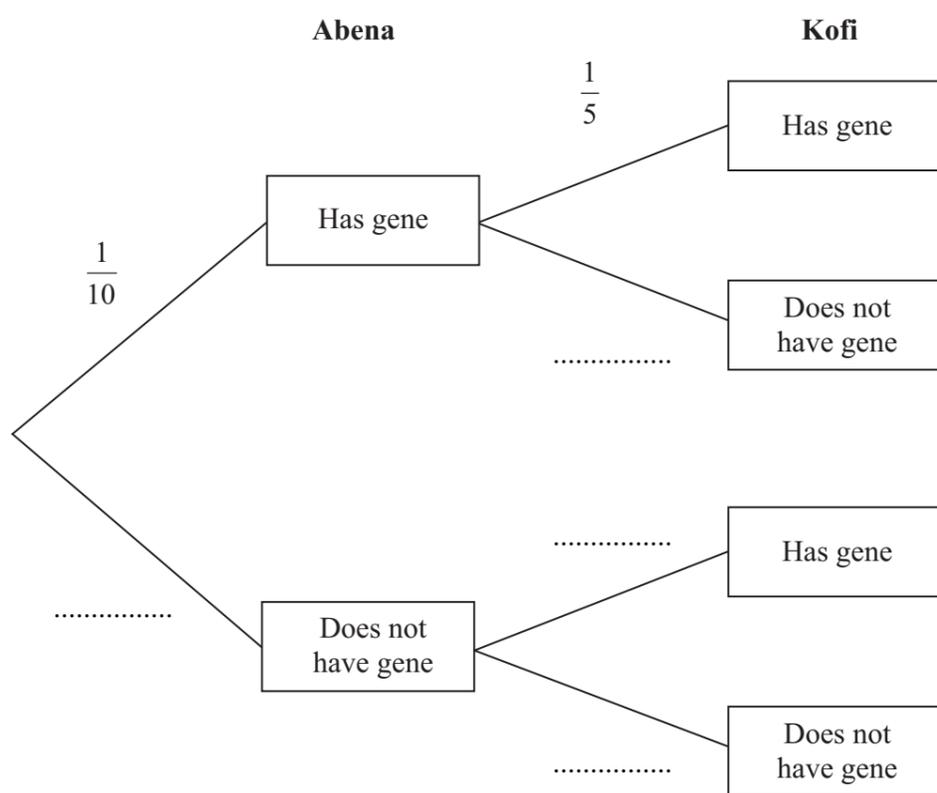
(2)

(b) Work out the probability that both Abena and Kofi have a sickle cell gene.

....

(2)

Abena and Kofi want to have 4 children together.

The probability that any child born to them will have sickle cell disease is $\dfrac{1}{4}$

(c) Work out the probability that at most one of their 4 children will have sickle cell disease.

[You may use $(p+q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$]

....
(4)

Q7

(Total 8 marks)

TOTAL FOR SECTION B: 65 MARKS

TOTAL FOR PAPER: 100 MARKS

END

BLANK PAGE

BLANK PAGE

BLANK PAGE