Please write clearly in block capitals.

Centre number ___________________________ Candidate number ___________________________

Surname ___________________________

Forename(s) ___________________________

Candidate signature ___________________________

GCSE STATISTICS

Higher tier Paper 2

Date of Exam Morning Time allowed: 1 hour 45 minutes

Materials

For this paper you must have:

• a calculator
• mathematical instruments.

Instructions

• Use black ink or black ball-point pen. Draw diagrams in pencil.
• Fill in the boxes at the top of the page.
• Answer all questions.
• You must answer the questions in the spaces provided. Do not write outside the box around each page or on blank pages.
• Do all rough work in this book. Cross out any work you do not want to be marked.

Information

• The marks for questions are shown in brackets.
• The maximum mark for this paper is 80.
• You may ask for more answer paper and graph paper. These must be tagged securely to this answer booklet.
<table>
<thead>
<tr>
<th>Question</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Paul gives a survey to every 5th student on the school registers. Circle the name for this type of sampling.</td>
</tr>
<tr>
<td></td>
<td>Random Stratified Systematic Quota</td>
</tr>
<tr>
<td>2</td>
<td>The mean of six numbers is 4 The first five of the numbers are 2 0 0 4 10 Circle the value which is the median of the six numbers.</td>
</tr>
<tr>
<td></td>
<td>0 2 3 5</td>
</tr>
<tr>
<td>3</td>
<td>Circle the name of the diagram that can be correctly used for grouped continuous data.</td>
</tr>
<tr>
<td></td>
<td>Bar chart Frequency polygon Pie chart Bar line chart</td>
</tr>
</tbody>
</table>
A doctor investigates how likely children are to have hay fever. She collects the following information from her patients.

<table>
<thead>
<tr>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{8} ) have hay fever</td>
<td>90 have hay fever</td>
</tr>
<tr>
<td></td>
<td>270 do not have hay fever</td>
</tr>
</tbody>
</table>

How many times more likely is hay fever in boys compared with in girls? Circle the answer.

[1 mark]

0.5  2  2.67  32
Jenny is doing a survey on people (tenants) who rent flats. She uses two rental companies ‘Letsmove’ and ‘Supaflat’. Her hypothesis is, “‘Letsmove’ tenants make fewer complaints than ‘Supaflat’ tenants.”

5 (a) Give two reasons why Jenny should take a sample and not ask every tenant. [2 marks]

Reason 1

Reason 2

5 (b) Jenny decides to take a sample using stratification.

5 (b) (i) Name one category which she could use to stratify her sample. [1 mark]

Answer

5 (b) (ii) Give a reason for your answer. [1 mark]
5 (c) Write a question which Jenny could use to find out the number of complaints a tenant had made.
Include a response section. [4 marks]

5 (d) Jenny is considering collecting the data using either telephone interviews, door to door interviews or an internet survey.

Which method would you choose from her list?

Data collection method ________________________

Give one advantage of your method over the other two methods. [1 mark]

______________________________
6 (a) Bag A contains 8 red and 7 blue counters.
Bag B contains 12 red and 10 blue counters.
A bag is chosen at random.
A counter is taken at random from the chosen bag.
Work out the probability that it is red.

Answer ________________

6 (b) Bag C contains only green and yellow counters.

- \[ P \text{ (green)} = \frac{3}{4} \]
- There are more than 20 but fewer than 30 counters in the bag.

Work out a possible value for the number of yellow counters there could be in the bag.

Answer ________________
A company makes a metal alloy by combining three metals, A, B and C, in the ratio 19 : 4 : 2.

The table shows the index numbers for the cost of each metal in 2016 using 2011 as the base year.

<table>
<thead>
<tr>
<th>Metal</th>
<th>Weighting</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>19</td>
<td>84.9</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>93.5</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>81.2</td>
</tr>
</tbody>
</table>

(a) Calculate a weighted index number for the combined cost of the metals in the alloy.

(b) The company claims that the combined cost of the metals in the alloy has fallen by over 15% between 2011 and 2016.

Is the company correct?
Tick a box.

Yes  No

Explain your answer.
A population pyramid is drawn to show the percentages of the UK population by age and gender in 2011.

The data for females and for some of the male ages has already been drawn.

8 (a) Use the table to complete the population pyramid for males.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Percentage of population (males)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 9</td>
<td>6.0</td>
</tr>
<tr>
<td>10 - 19</td>
<td>6.2</td>
</tr>
<tr>
<td>20 - 29</td>
<td>6.8</td>
</tr>
<tr>
<td>30 - 39</td>
<td>6.6</td>
</tr>
<tr>
<td>40 - 49</td>
<td>7.2</td>
</tr>
<tr>
<td>50 - 59</td>
<td>6.0</td>
</tr>
<tr>
<td>60 - 69</td>
<td>5.3</td>
</tr>
<tr>
<td>70 - 79</td>
<td>3.3</td>
</tr>
<tr>
<td>80+</td>
<td>1.2</td>
</tr>
</tbody>
</table>

[Source: Office for National Statistics]
8 (b) What percentage of the UK population are between the ages of 20 and 39? [2 marks]

Answer

8 (c) In 2011, the number of males aged 80 and over was 760 000

Calculate the number of males aged 10 – 19 years. [2 marks]

Answer

Turn over for the next question
Jane and Phil are studying house prices to compare Cumbria and Cornwall. They are going to send their findings to a local newspaper in Cumbria. Their hypothesis is 'house prices in Cornwall are more expensive than house prices in Cumbria.' They collect their data from a website which gives the house prices for all houses for sale in each area. They sort each list into price order and then collect their samples.

9 (a) Jane uses the first 30 house prices from each area.

What is the name of this sampling method? [1 mark]

9 (b) State one reason why this method will not produce a sample which is representative of the house prices in each area. [1 mark]

9 (c) Phil decides to use a different method to collect his sample.

Describe one method that Phil could use to collect a sample of 30 which is likely to be more representative of the house prices in each area. You should include the name of your sampling method, and a reason why a sample using this method is likely to be more representative. [4 marks]
Phil draws these boxplots.

9 (d) Give reasons why these diagrams.
   - are appropriate to comment on their hypothesis
   - support their hypothesis.

[3 marks]

9 (e) Phil says,
      “House prices are more varied in Cornwall.”

Is Phil correct?
Give a reason for your answer.

[1 mark]
Jane calculates the mean and range for each of her two sets of data.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumbria</td>
<td>£74 300</td>
<td>£48 500</td>
</tr>
<tr>
<td>Cornwall</td>
<td>£64 800</td>
<td>£50 000</td>
</tr>
</tbody>
</table>

9 (f) Write down **two** different interpretations that Jane could make using these values. Give **one** reason for **each** interpretation, write your answers so they can be understood by the readers of the local newspaper.

[4 marks]

9 (g) Jane decides to develop her study to include the number of bedrooms each house has. State **one** other variable that she could include to develop her study.

[1 mark]
It is claimed that feeding tomato plants with a new plant food, ‘Growfast’, will increase the number of tomatoes the plants produce. An experiment is to be set up to test this claim. Here is a list of variables that may be connected to the experiment.

A – How often the plant is watered.
B – The number of tomatoes a plant produces.
C – How much sunlight the plant gets.
D – The colour of the pot the tomatoes grow in.
E – Use of the plant food ‘Growfast’.

For this experiment

(a) circle the explanatory variable,

(b) circle the response variable.

(c) In the experiment, 50 tomato plants are fed ‘Growfast’ and 50 tomato plants are put into a control group.

(i) Explain the purpose of using a control group in this context.

(ii) Identify one possible extraneous variable from the list A, B, C, D, E and state how this might be controlled.
Mike asked a sample of 100 people which of the following countries they had visited.

Spain  France  Italy

- 6 had visited Spain and France and Italy
- 18 had visited France and Italy
- 36 had visited Spain and France
- 4 had visited Spain and Italy but not France
- 74 had visited Spain, 56 had visited France and 24 had visited Italy

11 (a) Use the data to complete the Venn diagram.

11 (b) How many people had visited none of the three countries?

Answer: ________________________________
11 (c) One of the 100 people is chosen at random.

Find the probability that the chosen person has visited

11 (c) (i) Spain or France but not Italy. [2 marks]

Answer ________________________________

11 (c) (ii) Spain given that they had visited Italy. [2 marks]

Answer ________________________________

11 (c) (iii) all three countries, given that they had visited at least two. [2 marks]

Answer ________________________________

Turn over for the next question
The histogram shows information about the time (minutes) that a sample of 175 passengers had to wait for a bus.

12 (a) Estimate the probability that a passenger, chosen at random from those who had to wait 6 minutes or more for a bus, actually had to wait 12 minutes or more.

[3 marks]
12 (b) Give one reason why the median may be a better measure of average than the mean in this situation. [1 mark]
13 A local firm makes yoghurt.

The graph shows the sales of yoghurt, in tonnes, from Quarter 1 of 2013 to Quarter 4 of 2015.

The four-point moving averages are also shown on the graph.

13 (a) Draw the trend line. [1 mark]
13 (b) Calculate the mean seasonal variation for Quarter 1 for the three years. [3 marks]

Answer ________________________________ tonnes

13 (c) Use your answer to part (b) and the trend line to predict the sales of yoghurt in Quarter 1 of 2016 [2 marks]

Answer ________________________________ tonnes

13 (d) Discuss the accuracy of the prediction you made in part (c). [2 marks]

________________________________________

________________________________________

________________________________________

________________________________________
Each weekday morning, Jon drives to work. His journey includes going over a railway level crossing where on any given day there is a 25% chance he is delayed.

(a) Explain why the number of days in one working week that he is delayed at the level crossing follows a binomial distribution.

(b) Show that the probability, in one working week, that he is delayed exactly once is 0.40 to two decimal places.
15 Kirstie is estimating the population of fish in a lake. 
She catches some fish and marks them with an harmless dye. 
She then returns them to the lake. 
One week later she catches a smaller sample of 50 fish and sees that 6 of them are marked. 
She correctly estimates there are 1125 fish in the lake.

15 (a) How many fish did she originally mark?  [3 marks]

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

15 (b) State two assumptions Kirstie makes to ensure this process is valid. 
Evaluate one of these assumptions; stating clearly which one it is.  [3 marks]

Assumption 1 __________________________________________________________________________
________________________________________________________________________

Assumption 2 __________________________________________________________________________
________________________________________________________________________

Evaluation ____________________________________________________________________________
________________________________________________________________________

END OF QUESTIONS