

Paper Reference 1ST0/2H
Pearson Edexcel
Level 1/Level 2 GCSE (9–1)

Total Marks

Statistics

Paper 2

(Calculator)

Higher Tier

Tuesday 16 June 2020 – Morning

**Time: 1 hour 30 minutes plus your
additional time allowance.**

**In the boxes below, write your name,
centre number and candidate number.**

Surname					
Other names					
Centre Number					
Candidate Number					

Y62714A



Pearson

YOU MUST HAVE

Ruler, protractor, compasses, writing and drawing equipment, scientific calculator.

YOU WILL BE GIVEN

Data Book

Formulae Pages

Turn over

INSTRUCTIONS

Answer ALL questions.

Answer the questions in the spaces provided in this Question Paper or on the separate data sheets – there may be more space than you need.

Scientific calculators may be used.

You must show all your working out with your answer clearly identified at the end of your solution.

You must NOT write anything on the Formulae Pages. Anything you write on the Formulae Pages will gain NO credit.

Turn over

INFORMATION

The total mark for this paper is 80

The marks for EACH question are shown in brackets – use this as a guide as to how much time to spend on each question.

There may be spare copies of some data sheets.

ADVICE

Read each question carefully before you start to answer it.

Try to answer every question.

Check your answers if you have time at the end.

5

Answer ALL questions.

Write your answers in the spaces provided.

You must write down all the stages in your working.

Turn over

- 1. Look at the diagram for Question 1 in the Data Book.**

It shows a frequency polygon.

The number of children living in each of 30 houses in a village was recorded in 1918 and in 2018

The frequency polygon has been drawn for the number of children recorded in 1918

The table on the next page shows information about the number of children living in each of the 30 houses in 2018

Turn over

1. continued.

Number of children	Frequency
0	6
1	9
2	8
3	4
4	2
5	1

(continued on the next page)

Turn over

1. continued.

(a) On the grid, draw a frequency polygon for the information for 2018

(2 marks)

(b) Compare the distributions of the number of children in these 30 houses in 1918 and in 2018

(2 marks)

Answer lines continue on the next page.

Turn over

1. (b) continued.

(Total for Question 1 is 4 marks)

Turn over

- 2. Look at the diagram for Question 2 in the Data Book.**

It shows a scatter diagram.

Mike is investigating whether or not there is a relationship between the distance from the nearest car park to a restaurant and the annual profit made by the restaurant.

(continued on the next page)

2. continued.

**He finds out the distance, in metres,
from the nearest car park to a
restaurant and the annual profit, in £,
made by the restaurant.**

**He does this for twenty restaurants in
a city.**

**Mike drew the scatter diagram for this
information.**

It is shown in the Data Book.

(continued on the next page)

Turn over

2. continued.

(a) Explain why annual profit is the response variable for this scatter diagram.

(1 mark)

(continued on the next page)

Turn over

2. continued.

Mike's hypothesis is that, for these restaurants, the further the restaurant is from its nearest car park the less is its annual profit.

(b) Explain, giving a statistical reason, whether or not the scatter diagram supports Mike's hypothesis.

(2 marks)

Answer lines continue on the next page.

Turn over

2. (b) continued.

(continued on the next page)

Turn over

2. continued.

Mike wants to draw a line of best fit on the scatter diagram.

Using statistical software he obtains the following information about these restaurants.

Mean distance from the nearest car park	325 metres
Mean annual profit	£27 000
Intercept of the line of best fit on the Annual profit axis	40

(continued on the next page)

Turn over

2. continued.

(c) (i) Using the information, draw a line of best fit on the scatter diagram.

(2 marks)

(ii) Interpret the value of the intercept of the line of best fit on the Annual profit axis.

(1 mark)

(continued on the next page)

Turn over

2. continued.

Restaurant A and restaurant B are two other restaurants in the city.

Restaurant A is 250 metres from its nearest car park.

Restaurant B is 700 metres from its nearest car park.

Mike uses the scatter diagram to find an estimate for the annual profit of each of these restaurants.

(continued on the next page)

Turn over

2. continued.

(d) Explain which of these two estimates will be the more reliable estimate.

(2 marks)

(continued on the next page)

Turn over

2. continued.

Mike finds a positive correlation between the number of tables at a restaurant and its annual profit.

He concludes that as the number of tables increases this causes the annual profit to increase.

(continued on the next page)

Turn over

2. continued.

(e) Explain whether or not this conclusion is valid.

(1 mark)

Mike reads an article in a newspaper that says that restaurant profits for the top 100 restaurants had fallen from £345 million to £125 million in the past year.

(Source: telegraph.co.uk)

(continued on the next page)

Turn over

2. continued.

(f) Using the data in the newspaper article, calculate the percentage decrease in restaurant profits for the top 100 restaurants in the past year.

(1 mark)

_____ %

(Total for Question 2 is 10 marks)

Turn over

- 3. Look at the diagram for Questions 3(a) and 3(b) in the Data Book.**

It shows an incomplete histogram.

The Forestry Commission planted Field Maple trees and Silverleaf Maple trees in region A

They measured the heights of the trees after 14 years.

The incomplete histogram in the Data Book and incomplete grouped frequency table on the next page give information about the heights, in metres, of the Field Maple trees in region A

(continued on the next page)

Turn over

3. continued.

Height (h metres)	Frequency
$0.8 < h \leq 1.2$	2
$1.2 < h \leq 1.6$	4
$1.6 < h \leq 2.0$	14
$2.0 < h \leq 2.4$	24
$2.4 < h \leq 2.8$	10
$2.8 < h \leq 3.2$	
$3.2 < h \leq 3.6$	

(Source: data.gov.uk)

(continued on the next page)

Turn over

3. continued.

(a) Use the information in the histogram to complete the table.

There are two spaces to fill.

(2 marks)

(b) Use the information in the table to complete the histogram in the Data Book.

(2 marks)

(continued on the next page)

Turn over

3. continued.

**Look at the diagram for Question 3(c)
in the Data Book.**

It shows a different histogram.

**It gives information about the heights
of the Silverleaf Maple trees after
14 years.**

**(c) Identify and interpret the type of
skew shown in the histogram for
Silverleaf Maple trees.**

(2 marks)

**Answer lines are on the next
page.**

Turn over

3. (c) continued.

Look at the table for Question 3(d) in the Data Book.

It shows a grouped frequency table.

**The Forestry Commission also
planted Field Maple trees and
Silverleaf Maple trees in region B**

**The grouped frequency table gives
information about the heights of
these trees.**

(continued on the next page)

Turn over

3. continued.

The estimate of the mean for Field Maple trees is calculated to be 206.3 cm to 1 decimal place.

David thinks that the estimate of the means for Field Maple trees and for Silverleaf Maple trees suggests that Field Maple trees are taller than Silverleaf Maple trees.

(continued on the next page)

Turn over

3. continued.

(d) Is David correct?

You must show your working.

**Give one limitation of your
conclusion.**

(4 marks)

(Total for Question 3 is 10 marks)

Turn over

- 4. Matthew is collecting information about road safety in his town.**

He wants to collect information from drivers who live in his town.

Matthew plans to ask drivers at his local car park about their views on road safety.

He plans to collect a quota sample of 20 drivers from each of three age groups.

(continued on the next page)

Turn over

4. continued.

**Matthew's three age groups of drivers
are**

39 years old or younger

40 years old to 59 years old

60 years old or older

(continued on the next page)

Turn over

4. continued.

(a) Comment on whether Matthew's plans are appropriate.

(2 marks)

(continued on the next page)

Turn over

4. continued.

Matthew wants to know how many drivers, who live in his town, drive faster than the speed limit on the motorway.

Matthew knows that he will have to ask a sensitive question so he plans to use the random response technique to find out this information.

(continued on the next page)

Turn over

4. continued.

(b) Design a random response question that Matthew could use on a questionnaire in order to collect this information.

(3 marks)

Answer space continues on the next page.

Turn over

4. (b) continued.

(Total for Question 4 is 5 marks)

Turn over

- 5. A supermarket manager is planning the design of a new petrol station.**

He wants to use the results of spinning a coin to simulate on which side of a car the fuel filler cap is.

He assumes that the fuel filler cap is equally likely to be on the left of the car as it is to be on the right of the car.

He wants to use the simulation in order to predict how many cars will arrive at the petrol station before there are four consecutive cars with filler caps on the same side.

(continued on the next page)

Turn over

5. continued.

(a) Explain how he could use a coin to carry out the simulation.

(2 marks)

(continued on the next page)

Turn over

5. continued.

**Look at Table 1 and Table 2 for
Question 5(b) in the Data Book.**

**A new cafe is also opening at the
supermarket.**

**Three different types of hot drink –
tea, coffee and hot chocolate – will be
sold in the cafe.**

**Table 1 gives, for each type of hot
drink, the fraction of the total number
of hot drinks sold in other cafes of
this supermarket chain.**

(continued on the next page)

Turn over

5. continued.

The manager wants to predict how many of each type of hot drink the cafe will sell in an hour the first Saturday the cafe is open.

The manager expects that there will be 40 customers in the cafe that hour.

He assigns the numbers shown in Table 2 to each type of hot drink.

(continued on the next page)

Turn over

5. continued.

He will then use a calculator to generate 40 random numbers between 0 and 9 to simulate 40 customers.

(b) Comment on the suitability of the manager's plan.

(3 marks)

Answer lines continue on the next page.

Turn over

5. (b) continued.

(Total for Question 5 is 5 marks)

6. Look at the diagram for Question 6 in the Data Book.

It shows two population pyramids.

They show the percentage of males and the percentage of females in each age group in the UK and in India for 2016

Each pyramid is based on the total population of that country for 2016

(continued on the next page)

Turn over

6. continued.

(a) Compare the percentages of males and females in the UK who are 24 years old or younger with the percentages of males and females in India who are 24 years old or younger.

You should give two comparisons.

(2 marks)

Answer lines continue on the next page.

Turn over

6. (a) continued.

(continued on the next page)

Turn over

6. continued.

Look at the information and at the table for Question 6(b) in the Data Book.

The population of the UK in 2016 was 65 648 000 and the number of live births to women aged 15 to 49 was 775 300
(Source: ons.gov.uk)

(continued on the next page)

Turn over

6. continued.

**(b) Calculate the general fertility rate
for the UK in 2016**

**You may use the formula given in
the Data Book.**

(4 marks)

**Answer space continues on the
next page.**

Turn over

6. (b) continued.

(Total for Question 6 is 6 marks)

Turn over

- 7. Look at the diagram for Question 7(a) in the Data Book.**

It shows an incomplete Venn diagram.

The manager of a garage records optional extra features sold with each new car sold last month.

The optional extra features are air conditioning (set **A), bluetooth (set **B**) and sat nav (set **S**)**

(continued on the next page)

7. continued.

Last month,

80 cars were sold

**7 cars were sold with sat nav and
bluetooth and air conditioning**

**15 cars were sold with sat nav and
bluetooth only**

**21 cars were sold with sat nav and
air conditioning only**

**10 cars were sold with bluetooth and
air conditioning only**

8 cars were sold with sat nav only

**33 cars in total were sold with
bluetooth**

**44 cars in total were sold with
air conditioning.**

(continued on the next page)

Turn over

7. continued.

**(a) Complete the Venn diagram
using the information on the
previous page.**

(3 marks)

**Inge picks at random a car that was
sold last month.**

**She thinks that the probability that
the car has bluetooth given that it has
sat nav is greater than the probability
that the car has bluetooth given that
it has not got sat nav.**

(continued on the next page)

Turn over

7. continued.

(b) Is Inge correct?

You must show your working.

(5 marks)

**Answer space continues on the
next page.**

Turn over

7. (b) continued.

(continued on the next page)

Turn over

7. continued.

Look at the table for Question 7(c) in the Data Book.

Second-hand cars are also sold at the garage.

The manager records whether the second-hand cars have a full service history or not.

She also records whether each second-hand car has a breakdown or not in the first year after being sold.

(continued on the next page)

Turn over

7. continued.

The table in the Data Book gives information about the second-hand cars sold at the garage in 2018

Paul says that the relative risk of a second-hand car without a full service history having a breakdown in the first year after being sold, compared with a second-hand car with a full service history having a breakdown in the first year after being sold, is 2

(continued on the next page)

Turn over

7. continued.

(c) (i) Show that Paul is correct.

(2 marks)

(continued on the next page)

Turn over

7. (c) continued.

**(ii) Interpret a relative risk of 2
in this context.**

(1 mark)

(Total for Question 7 is 11 marks)

Turn over

- 8. Look at the diagram for Question 8(a) in the Data Book.**

It shows a scatter diagram.

The Spearman's rank correlation coefficient and the Pearson's product moment correlation coefficient are each going to be calculated for the data shown in the scatter diagram.

(continued on the next page)

8. continued.

**(a) How would you expect the
values of these two correlation
coefficients to compare?**

(2 marks)

(continued on the next page)

Turn over

8. continued.

Look at the table for Question 8(b) in the Data Book.

The Australian Government collected data on crop yields from fields.

The data collected were the yield of wheat grown in a field, in tonnes per hectare (t/ha), and the yield of other crops grown in the same field (t/ha) in a different year.

(continued on the next page)

Turn over

8. continued.

The table gives the Pearson's product moment correlation coefficient for the wheat yield and the barley yield and for the wheat yield and the oats yield. The table also gives the equation of the regression line in each case.

(continued on the next page)

8. continued.

(b) Compare the relationship between wheat yield and barley yield with the relationship between wheat yield and oats yield.

You should refer to both correlation coefficients and to the equations of both regression lines in your comparison.

(6 marks)

Answer lines continue on the next page.

8. (b) continued.

(continued on the next page)

Turn over

8. continued.

This year Louise has a field planted with wheat.

She wants to use the information in the table to decide whether to plant barley or to plant oats in the field next year.

She plans to plant the crop which is predicted to give the greatest yield.

(continued on the next page)

Turn over

8. continued.

Louise solves the equation

$$1.24x - 0.30 = 1.52x - 1.05$$

She gets the answer $\frac{75}{28}$

(c) (i) Show that the answer Louise gets is correct.

(2 marks)

Answer space continues on the next page.

Turn over

8. (c) (i) continued.

(continued on the next page)

Turn over

8. (c) continued.

(ii) Explain how Louise could use her answer to decide which of barley or oats to plant next year.

(2 marks)

(continued on the next page)

Turn over

8. (c) continued.

**(iii) Give a limitation of the data
that could affect the result of
Louise's decision.**

(1 mark)

(Total for Question 8 is 13 marks)

Turn over

9. A company manufactures tablets of medication **X**

The tablets have a target mass of medication **X** of **200 mg**

The company uses quality assurance to monitor the mass of medication **X** in each tablet.

Samples of the tablets are taken from the production line at regular intervals and the mean mass of medication **X** in the tablets in each sample is found.

(continued on the next page)

Turn over

9. continued.

The sample means should be normally distributed with a mean of 200 mg and a standard deviation of 2.5 mg

- (a) Find the upper action limit for the sample means for medication X (2 marks)**

Answer space continues on the next page.

Turn over

9. (a) continued.

_____ mg

(continued on the next page)

Turn over

9. continued.

The pharmacist in charge of monitoring the production of the tablets of medication **X** wants to set the upper action limit closer to the target mass of **200 mg**

(continued on the next page)

Turn over

9. continued.

(b) What effect would you expect this to have on the number of times the production process may need to be stopped?

(1 mark)

(continued on the next page)

Turn over

9. continued.

**Look at the diagrams for
Question 9(c) in the Data Book.
They show two control charts.**

**The company also manufactures
tablets of medication Y and the
company uses quality assurance to
monitor the mass of medication Y in
the tablets.**

**The control charts for the sample
means and for the sample ranges
of the mass of medication Y in the
tablets are shown in the Data Book.**

(continued on the next page)

Turn over

9. continued.

A sample of tablets of medication Y is taken.

The mean of the sample is 107 mg

The range of the sample is 39 mg

(c) Determine what action, if any, should be taken.

Justify your answer by referring to both the sample mean and the sample range.

(2 marks)

Answer lines are on the next page.

Turn over

9. (c) continued.

(Total for Question 9 is 5 marks)

Turn over

10. Look at the table for Question 10(a) in the Data Book.

It shows a grouped frequency table.

Kate is investigating the lengths of earthworms.

She wants to find out if the lengths of common earthworms can be modelled by a normal distribution.

She has found out that the greatest length of the common earthworm is 36 cm

(Source: www.nationalgeographic.com)

(continued on the next page)

Turn over

10. continued.

Kate plans to group the data that she will collect in the table shown in the Data Book.

(a) Comment on whether Kate's choice of class intervals in the grouped frequency table is appropriate.

(2 marks)

Answer lines continue on the next page.

Turn over

10. (a) continued.

(continued on the next page)

Turn over

10. continued.

Look at the table and at the information for Question 10(b) in the Data Book.

Bien is researching the lengths of time, in days, lived by flies.

He wants to know if the lengths of time lived by flies can be modelled by a normal distribution.

He has collected data for 80 flies.

(continued on the next page)

Turn over

10. continued.

The table in the Data Book gives information about the length of time lived by each fly.

Look at Bien's plan in the Data Book to test whether the lengths of time lived by flies can be modelled by a normal distribution.

(b) Comment on whether Bien's plan is appropriate.

(4 marks)

Answer lines are on the next two pages.

Turn over

10. (b) continued.

Turn over

10. (b) continued.

(Total for Question 10 is 6 marks)

**11. Look at the information for
Question 11 in the Data Book.**

**Bella collected the finish times,
X minutes, of the men's elite
wheelchair race at the 2018 London
Marathon.**

**She uses statistical software to
calculate the summary statistics as
shown in the Data Book.**

(continued on the next page)

Turn over

11. continued.

Calculate the skew for the distribution of the finish times and interpret this value in the context of Bella's data.

(5 marks)

Answer space and lines continue on the next two pages.

Turn over

11. continued.

Turn over

11. continued.

(Total for Question 11 is 5 marks)

TOTAL FOR PAPER IS 80 MARKS

END OF PAPER
