

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

Total Marks

Statistics
PAPER 2
Higher Tier
(Calculator)

Monday 19 June 2023 – Afternoon

Time: 1 hour 30 minutes

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

Surname					
Other names					
Centre Number					
Candidate Number					

YOU MUST HAVE

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

YOU WILL BE GIVEN

**Data Booklet
Formulae Pages**

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.

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 in case you need them.

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.

Turn over

Answer ALL questions.

Write your answers in the spaces provided.

You must write down all the stages in your working.

- 1. Grace asked a sample of 60 people in her town if they had ever visited France or Spain.**

17 people visited both France and Spain

23 people visited Spain only

33 people visited France

- (a) Draw a Venn diagram in the Data Booklet to represent this information.**

There is blank space on pages 4 and 17 in the Data Booklet.

(5 marks)

1. continued.

Grace says

- **more than half of the people in her sample have visited France**
- **therefore more than half of the people in her town have visited France**

(continued on the next page)

Turn over

1. continued.

**(b) Discuss the validity of each of
Grace's comments.**

(3 marks)

(Total for Question 1 is 8 marks)

Turn over

2. Look at the diagram for Question 2 in the Data Booklet.

It is a box plot.

Logan is investigating the heights of male adult giraffes and the heights of female adult giraffes.

He records the height, in metres, of each of a sample of male adult giraffes and the height, in metres, of each of a sample of female adult giraffes.

(continued on the next page)

2. continued.

He draws the box plot in the Data Booklet for the recorded heights of the male adult giraffes.

The table on the next page gives information about the recorded heights of the female adult giraffes.

(continued on the next page)

Turn over

2. continued.

Summary statistic	Height (metres)
Mean	4·8
Median	4·9
Minimum	3·9
Maximum	5·9
Lower quartile	4·2
Upper quartile	5·4

(continued on the next page)

Turn over

2. continued.

Logan makes the following two conclusions.

- 1. Male adult giraffes are generally taller than female adult giraffes.**
- 2. The heights of the female adult giraffes are more consistent than the heights of the male adult giraffes.**

(continued on the next page)

Turn over

2. continued.

Assess Logan's two conclusions.

You should show clearly the values of any statistics you use in your answer.

(5 marks)

Answer lines are on the next two pages.

Turn over

2. continued.

Turn over

2. continued.

(Total for Question 2 is 5 marks)

3. Look at the diagram for Question 3 in the Data Booklet.

It is a cumulative frequency step polygon showing information about the number of goals scored in each of 28 matches played by the German women's national football team.

(continued on the next page)

3. continued.

(a) Give a reason why a cumulative frequency step polygon is used to represent this information rather than a cumulative frequency curve.

(1 mark)

(continued on the next page)

Turn over

3. continued.

(b) Find the mode of the number of goals scored.

(1 mark)

(continued on the next page)

Turn over

3. continued.

(c) Find the number of these matches where

**(i) exactly 6 goals were scored,
(1 mark)**

(continued on the next page)

Turn over

3. (c) continued.

**(ii) more than 6 goals were
scored.**

(2 marks)

(continued on the next page)

Turn over

3. continued.

In 24 matches fewer than n goals were scored.

**(d) Find the value of n
(1 mark)**

(continued on the next page)

Turn over

3. continued.

**Klara tries to calculate the
interquartile range of the number of
goals scored.**

She gets an answer of 14

**(e) Explain how you know that her
answer is incorrect.**

(1 mark)

(Total for Question 3 is 7 marks)

Turn over

- 4. Maya works in a factory that produces hammers.**

Today she is going to take a sample of the hammers produced and check them for defects.

- (a) Which one of these statements best describes the population for Maya's sample?**

This question is multiple choice.

Write the letter of your chosen answer in the box provided on the following page.

Turn over

4. (a) continued.

A 50 employees at the factory

B 100 hammers produced today

C all of the employees at the factory

D all of the hammers produced today

Answer

(1 mark)

(continued on the next page)

Turn over

4. continued.

Maya wants to avoid bias in her sample.

She obtains a list of the unique identification numbers for each hammer.

Maya is deciding between using quota sampling and systematic sampling.

(continued on the next page)

4. continued.

**(b) For each sampling method,
assess whether or not it would
be an appropriate method to use
for Maya's sample.**

**(i) quota sampling
(2 marks)**

(continued on the next page)

Turn over

4. (b) continued.

(ii) systematic sampling

(2 marks)

(Total for Question 4 is 5 marks)

5. Look at the diagram for Question 5 in the Data Booklet.

It shows two spinners that Roslyn makes and two tables.

To test if each spinner is fair, she spins each spinner 60 times and records the side it lands on.

The tables in the Data Booklet show her results.

Roslyn believes one of the spinners is biased.

(continued on the next page)

Turn over

5. continued.

Discuss Roslyn's belief.

**You should comment on the
outcomes for both spinners in your
discussion.**

(4 marks)

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

Turn over

5. continued.

(Total for Question 5 is 4 marks)

6. The table shows information about the times, in minutes, taken for 60 people to complete a 5 km race.

Time (t minutes)	Frequency
$16 < t \leq 20$	5
$20 < t \leq 24$	25
$24 < t \leq 28$	15
$28 < t \leq 32$	10
$32 < t \leq 36$	5

(continued on the next page)

Turn over

6. continued.

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

It is an incomplete frequency polygon.

The incomplete frequency polygon is drawn to represent the information in the table.

(a) Work out the number of people who completed the race in 24 minutes or less.

(1 mark)

Answer space continues on the next page.

Turn over

6. (a) continued.

(continued on the next page)

6. continued.

**(b) Complete the frequency polygon
in the Data Booklet.**

(2 marks)

(continued on the next page)

6. continued.

One of the people who completed the race is selected at random.

(c) Find the probability that this person completed the race in more than 20 minutes.

(1 mark)

6. continued.

(d) Explain how the shape of this distribution of times can be used to compare the mean time with the median time.

(2 marks)

(continued on the next page)

Turn over

6. continued.

The 1st decile for these results is 21 minutes and the 9th decile for these results is 30 minutes.

(e) Work out an estimate for the 1st decile to 9th decile interdecile range.

(1 mark)

Answer space continues on the next page.

6. (e) continued.

_____ minutes

(continued on the next page)

Turn over

6. continued.

- (f) Find the class interval that contains the 65th percentile.
(1 mark)**

(Total for Question 6 is 8 marks)

Turn over

- 7. Look at the table for Question 7 in the Data Booklet.**

The table shows information about the death rate in England and Wales in 2000

- (a) Using the formula below, calculate the crude death rate for England and Wales in 2000 (2 marks)**

Answer space is on the next page.

crude death rate =

$$\frac{\text{number of deaths} \times 1000}{\text{total population}}$$

Turn over

7. (a) continued.

(continued on the next page)

Turn over

7. continued.

**In 2000, the crude birth rate in
England and Wales was 11·6**

**In 2019, the crude birth rate in
England and Wales was 10·8**

**A newspaper prints the following
headline for an article.**

**“36 000 more births in England and
Wales in 2019 than in 2000”**

(continued on the next page)

Turn over

7. continued.

(b) Discuss whether or not this headline could be true given that the crude birth rate in 2019 is less than the crude birth rate in 2000

(2 marks)

(Total for Question 7 is 4 marks)

Turn over

8. A scientist is carrying out an experiment to test the effect of a vitamin on memory.

She plans to use matched pairs.

(a) Explain what is meant by matched pairs in an experiment.

(2 marks)

(continued on the next page)

Turn over

8. continued.

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

Jamal is one of the participants in the experiment.

He takes four tests as part of the experiment.

Each test is given a different weighting.

The table in the Data Booklet shows the weighting and the score for each test.

(continued on the next page)

Turn over

8. continued.

(b) Calculate the weighted mean score for Jamal's four tests.

(3 marks)

Answer space continues on the next page.

8. (b) continued.

(Total for Question 8 is 5 marks)

Turn over

- 9. Look at the table for Question 9 in the Data Booklet.**

In a survey, children were asked to state their main source of news.

They were also asked how interested they were in the news.

The table in the Data Booklet shows the results of the survey for those stating Newspapers as their main source of news.

For those who said Radio is their main source of news, 65 responded ‘Very’ as their level of interest.

(continued on the next page)

Turn over

9. continued.

Comparative pie charts are drawn to represent the responses for Newspapers and for Radio.

(a) Compare the area of the sector for 'Very' in the pie chart for Newspapers with the area of the sector for 'Very' in the pie chart for Radio.

Give a reason for your answer.

(2 marks)

Answer lines continue on the next page.

9. (a) continued.

(continued on the next page)

9. continued.

The radius for the pie chart for Newspapers is 3 cm

The radius for the pie chart for Radio is 4.7 cm

(b) Calculate the size of the angle for the sector for 'Very' in the pie chart for Radio.

(3 marks)

Answer space continues on the next page.

Turn over

9. (b) continued.

○

(Total for Question 9 is 5 marks)

Turn over

10. Look at the diagram for Question 10 in the Data Booklet.

It shows an incomplete quality assurance report.

A company produces bolts with a diameter 16.4 mm

Tamika takes random samples of 4 bolts at regular intervals from the production line and the mean diameter of each sample is calculated.

Her incomplete quality assurance report is shown in the Data Booklet.

(continued on the next page)

Turn over

10. continued.

Complete the missing information in Tamika's quality assurance report by filling in the table and completing the graph in the Data Booklet.

(Total for Question 10 is 5 marks)

- 11. The heights of a group of seven-year-old boys have a mean of 121.7 cm and a standard deviation of 5.3 cm**

Daniel is a seven-year-old boy with a standardised score for height of 0

- (a) Write down Daniel's height.
(1 mark)**

_____ cm

(continued on the next page)

Turn over

11. continued.

**Syed and Timur are both
seven-year-old boys.**

**Syed's standardised score for height
is 1.4**

**Timur's standardised score for height
is -1.6**

Syed is taller than Timur.

(b) How much taller?

(3 marks)

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

Turn over

11. (b) continued.

_____ cm

(continued on the next page)

Turn over

11. continued.

Tarik takes a sample of 4 boys from the group of seven-year-old boys.

He wants to calculate the standardised score for the

SAMPLE MEAN of their heights.

He uses 121·7 cm for the mean and 5·3 cm for the standard deviation.

(continued on the next page)

Turn over

11. continued.

(c) Discuss whether or not it is appropriate to use

**(i) 121.7 cm as the mean
in the calculation of the
standardised score,
(2 marks)**

(continued on the next page)

Turn over

11. (c) continued.

**(ii) 5.3 cm as the standard
deviation in the calculation
of the standardised score.**

(2 marks)

(Total for Question 11 is 8 marks)

Turn over

12. Look at the table for Question 12 in the Data Booklet.

A bag contains red counters, blue counters and white counters only.

The table shows the proportion of each colour of counter in the bag.

In an experiment, Salome selects one of the counters from the bag and records its colour.

She then puts the counter back into the bag.

(continued on the next page)

12. continued.

She repeats these steps 4 more times, so that she makes a total of 5 selections.

Salome says that the probability she selects exactly 1 red counter is more than the probability that she selects exactly 2 blue counters.

(continued on the next page)

12. continued.

- (a) Use binomial distributions to determine if Salome is correct. You should show your working below and on the next two pages. (6 marks)**

12. (a) continued.

Turn over

12. (a) continued.

(continued on the next page)

12. continued.

(b) By considering the conditions that make a binomial distribution a suitable model, assess the appropriateness of using a binomial distribution to model the number of red counters that Salome selects.

(3 marks)

Answer lines continue on the next page.

Turn over

12. (b) continued.

(continued on the next page)

12. continued.

In a different experiment, Malena selects 1 of the counters from the bag and records its colour.

If the counter is red, she stops selecting.

Otherwise she puts the counter back in the bag and selects again.

She repeats this process until a red counter is selected.

Malena records the number of counters selected until she selects a red counter.

(continued on the next page)

Turn over

12. continued.

(c) Assess the use of a binomial distribution to model the number of counters selected by Malena.

(1 mark)

(Total for Question 12 is 10 marks)

Turn over

13. Indre is writing a report on the relationship between the income per person, in US dollars, and the life expectancy, in years, for different countries.

She finds data from the internet for 175 countries.

(continued on the next page)

13. continued.

(a) Give a reason why it is important for Indre to acknowledge the source of the data she uses in her report.

(1 mark)

(continued on the next page)

Turn over

13. continued.

Indre calculates Pearson's product moment correlation coefficient (PMCC) between income per person and life expectancy and she obtains a value of 0.72

**(b) Give an interpretation of this value in context.
(1 mark)**

(continued on the next page)

Turn over

13. continued.

**Look at the diagram for
Question 13(c) and 13(d) in the
Data Booklet.**

**Indre uses statistical software
to draw a scatter diagram for the
information collected as shown in the
Data Booklet.**

**She then uses the software to
calculate the value of Spearman's
rank correlation coefficient for this
data.**

(continued on the next page)

Turn over

13. continued.

(c) Describe how you would expect the value of Spearman's rank correlation coefficient for this data to compare with the value 0.72

Give a reason for your answer.

(2 marks)

(continued on the next page)

Turn over

13. continued.

**For the 175 countries, Indre
calculates Pearson's product moment
correlation coefficient between
fertility rate and life expectancy and
obtains a value of -0.80**

(continued on the next page)

13. continued.

(d) Compare the strength of the correlation between fertility rate and life expectancy with the strength of the correlation between income per person and life expectancy.

Give a reason for your answer.

(2 marks)

Answer lines continue on the next page.

Turn over

13. (d) continued.

(Total for Question 13 is 6 marks)

TOTAL FOR PAPER IS 80 MARKS

END OF PAPER

Sources

Question 6

(Source adapted from: www.parkrun.org.uk/southwark/results)

Question 11

(Source: www.who.int/growthref/who2007_height_for_age/en/)

Question 13(b)

(Source: <https://www.gapminder.org/data/>)

Question 13(d)

(Source: www.gapminder.org/data/)