Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number	Candidate Number

Time 2 hours

Paper reference **9ST0/02**

Statistics

Advanced
PAPER 2: Statistical Inference

You must have:
Statistical formulae and tables booklet
Calculator

Total Marks

Candidates may use any calculator allowed by Pearson regulations.
Calculators must not have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Unless otherwise stated, inexact answers should be given to three significant figures.
- Unless otherwise stated, statistical tests should be carried out at the 5% significance level.

Information

- A booklet 'Statistical formulae and tables' is provided.
- There are 7 questions in this question paper. 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.*

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.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.
- Good luck with your examination.

Turn over ▶

P66656A

1/1/1/1/1/1/1/1/

P 6 6 6 5 6 A 0 1 2 4

Pearson

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

1 Ramon is conducting a research project.

He has taken a random sample of 16 pairs of values for (x, y).

He has then calculated, from his data, a sample Pearson's product moment correlation coefficient of 0.716

Ramon believes that this calculated figure indicates that the population Pearson's product moment correlation coefficient is significantly different from zero.

(a) Conduct a hypothesis test to investigate Ramon's belief.

(4)

Question 1 continued

Ramon has also produced a scatter graph from his data, shown in **Figure 1**.

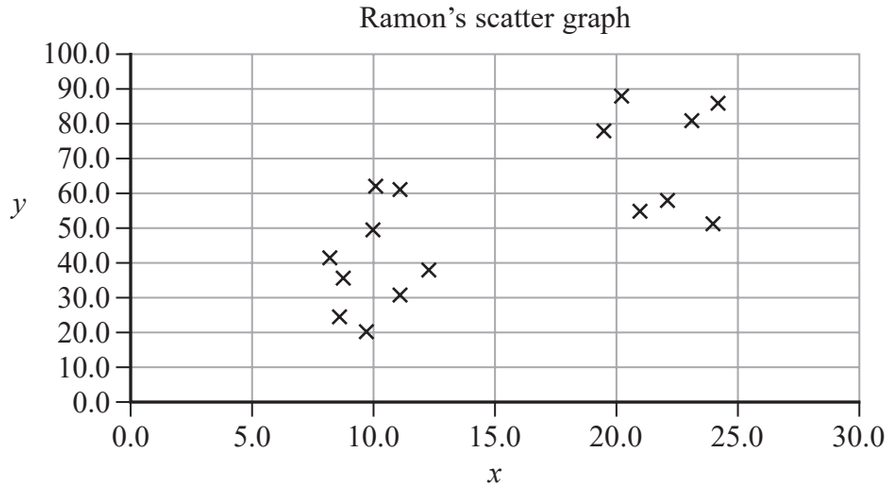


Figure 1

(b) With reference to the scatter graph, explain why the conclusion made from the hypothesis test conducted in (a) may not be valid.

(2)

(Total for Question 1 is 6 marks)

2 Mohamed grows courgettes on his farm.

He knows that the weights, in grams, of courgettes grown on other nearby farms are distributed $N(175, 22.0^2)$.

Mohamed decides to compare the weight of courgettes grown on his farm with the weight of courgettes grown on other nearby farms.

Mohamed takes a random sample of 16 courgettes grown on his farm. The mean weight of courgettes in this random sample is 182.9 grams.

(a) Calculate a 95% confidence interval for the mean weight of courgettes grown on Mohamed's farm.

You may assume that the standard deviation of weights of courgettes grown on Mohamed's farm is the same as that for courgettes grown on other nearby farms.

(3)

(b) Explain why the confidence interval you calculated in (a) does **not** provide significant evidence that courgettes grown on Mohamed's farm have a mean weight different from 175 grams.

(1)

Question 2 continued

Mohamed decides to take another, larger, random sample of courgettes grown on his farm.

(c) Assuming that his sample mean will remain 182.9 grams, calculate how large Mohamed's random sample should be in order to provide significant evidence that his courgettes have a mean weight different from 175 grams.

(3)

(Total for Question 2 is 7 marks)

3 Stewart is planning a survey of 85 randomly sampled business leaders in his industry. He believes that **most** business leaders in his industry would support their business adopting a four-day working week.

Stewart decides to include in his survey the question:

Would you support your business adopting a four-day working week?

Yes		**No**	

(a) Making any necessary assumptions, state a probability distribution that Stewart could use to model the number of **yes** responses.

(1)

Stewart wants to conduct a hypothesis test to investigate his belief.

(b) For Stewart's hypothesis test,

(i) state the **hypotheses** that he should use,

(1)

(ii) find the **critical region**.

(2)

Question 3 continued

(c) State the probability of a Type I error in the hypothesis test described in (b).

(1)

Debra believes that two-thirds of business leaders in Stewart's industry would support their business adopting a four-day working week.

(d) Assuming that Debra's belief is correct, calculate the power of the hypothesis test described in (b).

(3)

(Total for Question 3 is 8 marks)

4 An English lecturer, Eleanor, is interested in the relationship between vocabulary scores and reading habits. She decides to undertake a study to investigate this.

Eleanor asks all of her students to take a vocabulary test. She also asks them about their reading habits.

The outcome of the vocabulary test is a score. Students self-select a reading habit descriptor from a list of several such descriptors.

Eleanor decides to compare the scores of two groups of students who self-select the reading habit descriptors

- 'read **widely**', or
- 'read **fairly widely**'.

She takes independent random samples from each of these groups.

Eleanor summarises the scores in a table.

	Score	
	Students who self-select 'read **widely**'	Students who self-select 'read **fairly widely**'
Sample mean	32.2	29.4
Sample standard deviation	5.11	4.18
Sample size	15	15

(a) Making any necessary assumptions, investigate whether students who self-select 'read **widely**' have a significantly higher average score than students who self-select 'read **fairly widely**'.

(6)

Question 4 continued

(b) Give **two** criticisms of Eleanor's study.

(2)

(Total for Question 4 is 8 marks)

5 Diamond League is an athletics competition that takes place every year in 14 locations, including Monaco and London.

Tessa believes that athletic performances in different locations vary due to a range of factors including the weather and the stadium design.

She collected Diamond League race times for Monaco and London. The race times, in minutes and seconds, of all athletes finishing the Men's 800 metres (2019) are given in **Figure 2**.

Diamond League race times for the Men's 800 metres (2019)	
Monaco	**London**
1 m 41.89 s	1 m 43.14 s
1 m 42.54 s	1 m 43.48 s
1 m 43.62 s	1 m 43.74 s
1 m 43.70 s	1 m 43.90 s
1 m 43.83 s	1 m 44.42 s
1 m 44.40 s	1 m 44.52 s
1 m 44.69 s	1 m 44.61 s
1 m 45.08 s	1 m 44.75 s
1 m 45.43 s	1 m 44.97 s
1 m 45.78 s	1 m 45.03 s
	1 m 45.10 s

[Data source: *https://monaco.diamondleague.com/, https://london.diamondleague.com/*]

Figure 2

(a) Use the race times given in **Figure 2** to test whether the average race time for the Men's 800 metres in Monaco is different from that for the Men's 800 metres in London.

Assume that race times for the Men's 800 metres are **not** normally distributed.
Make any necessary assumptions regarding independence.

(8)

Question 5 continued

Question 5 continued

Tessa later noticed that there was an athlete who raced in both locations in the Men's 800 metres (2019).

Adam Kszczot achieved the time of 1 m 44.69 s in Monaco and 1 m 44.61 s in London.

(b) Comment, in light of this information about Adam, on whether the conditions for the test conducted in (a) are met.

(1)

(c) Suggest **two** different improvements to Tessa's sampling method.

(2)

(Total for Question 5 is 11 marks)

6 A research team conducted a study into the weights of a random sample of 5571 new-born lambs.

The mean weight of the lambs in the sample was 5.01 kg and the standard deviation of the weights of the lambs in the sample was 1.30 kg.

The research team wanted to investigate whether the distribution of the weights of the sample of lambs was consistent with the normal distribution.

The weights of the sample of lambs are summarised in **Figure 3**.

Also included in **Figure 3** are some expected frequencies calculated using a normal distribution based on the sample mean and standard deviation.

Weight range, W (kg)	**Observed frequency**	**Expected frequency N(5.01, 1.30^2)**	
$W < 1.5$	15	19.31	
$1.5 \leqslant W < 2.0$	32	38.04	
$2.0 \leqslant W < 3.0$	271	282.66	
$3.0 \leqslant W < 4.0$	865	877.81	
$4.0 \leqslant W < 5.0$	1600	1550.57	
$5.0 \leqslant W < 6.0$	1575	1559.33	
$6.0 \leqslant W < 7.0$	884	892.78	
$7.0 \leqslant W < 8.0$	271	290.75	
$8.0 \leqslant W < 9.0$	51		
$W \geqslant 9.0$	7		

Figure 3

(a) Complete the column of expected frequencies by calculating the **two** missing frequencies.

(2)

Question 6 continued

(b) Investigate whether the distribution of the weights of the sample of lambs is consistent with the normal distribution.

(8)

Question 6 continued

The lambs in the study belonged to **four** different breeds.

The research team conducted a one-factor ANOVA to compare the weights of lambs of the four different breeds.

(c) State the hypotheses to be used in the comparison.

(1)

Question 6 continued

The partially completed one-factor ANOVA table is given in **Figure 4**.

Source of variation	Sum of squares	Degrees of freedom	Mean square	F ratio
Between breeds	578.521			
Within breeds	148147.218			
Total	148725.739			

Figure 4

(d) Complete **Figure 4**.

(3)

Question 6 continued

(e) Use your one-factor ANOVA table to complete this ANOVA test for the comparison of mean weights of lambs for the **four** different breeds.

NOTE: If you are using Table 7 in the Statistical formulae and tables booklet to obtain the critical value, when the degrees of freedom are large you should use the critical value for ∞.

(3)

(f) State **two** assumptions needed to make your conclusion in part (e) valid.

(2)

(Total for Question 6 is 19 marks)

7 Trains arrive into stations either early, on time or late. An early arrival is recorded as a negative delay. An on-time arrival is recorded as zero delay. A late arrival is recorded as a positive delay.

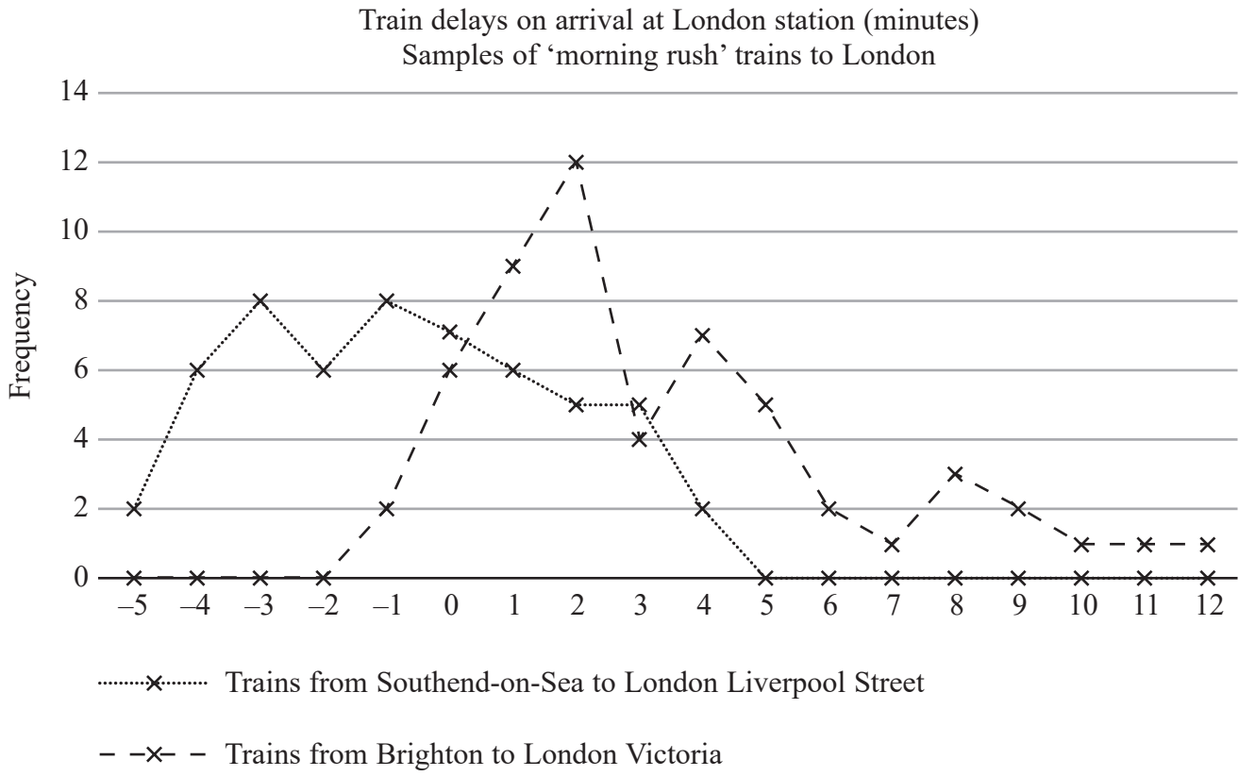
Richard has collected samples of train delays on two routes, for those 'morning rush' trains scheduled to depart between 0631 and 0830 on weekdays in summer 2019. These two routes will be the only routes considered in this question. The routes are

- Southend-on-Sea to London Liverpool Street,
- Brighton to London Victoria.

The total timetabled journey time on both routes is between 1 hour and 1 hour 5 minutes.

You may assume that Richard's samples are random.

The delays, to the nearest minute, are plotted on the frequency polygons.



[Data source: *recenttraintimes.co.uk*]

Richard has calculated these sample statistics.

Delay in minutes	Trains from Southend-on-Sea to London Liverpool Street	Trains from Brighton to London Victoria
mean	–0.709	3.39
standard deviation	2.45	3.10
sample size	55	56

Figure 5

Question 7 continued

(a) State **two** reasons why the mean sampled delays may be modelled as

$$\mathrm{N}\left(\mu_S, \frac{2.45^2}{55}\right) \text{ and } \mathrm{N}\left(\mu_B, \frac{3.10^2}{56}\right)$$

where μ_S and μ_B are the mean delays on all 'morning rush' trains from Southend-on-Sea and Brighton respectively.

(2)

(b) Conduct a hypothesis test using Richard's data to investigate whether the mean delay on 'morning rush' trains from Southend-on-Sea is less than zero (on average trains arrive early).

(6)

Question 7 continued

Richard decides to conduct a test to investigate whether the mean delay on 'morning rush' trains from Brighton is more **than 3 minutes greater** than the mean delay on 'morning rush' trains from Southend-on-Sea.

He calculates the following test value:

$$\frac{[3.39 - (-0.709)] - [3]}{\sqrt{\frac{2.45^2}{55} + \frac{3.10^2}{56}}} = 2.074$$

(c) Write down

- the hypotheses of Richard's test
- the critical region of Richard's test
- whether Richard's test value of 2.074 provides significant evidence against H_0.

(3)

Question 7 continued

(d) Find the p-value corresponding to Richard's test value of 2.074

Explain whether, in this case, there is an advantage to using the p-value rather than the critical region of the test.

(2)

In a further investigation, Richard noted that, during 44 weekdays in July and August, there were 8 days when at least one 'morning rush' train from Brighton was cancelled, but there were only 4 days when at least one 'morning rush' train from Southend-on-Sea was cancelled.

(e) Does this investigation provide significant evidence that there are more likely to be days when at least one 'morning rush' train is cancelled from Brighton than from Southend-on-Sea?

(6)

Question 7 continued

(f) State **two** statistical assumptions you made in (e).

(2)

(Total for Question 7 is 21 marks)

TOTAL FOR PAPER IS 80 MARKS

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

BLANK PAGE

BLANK PAGE