

Please check the examination details below before entering your candidate information

Candidate surname

Other names

**Pearson Edexcel  
Level 3 GCE**

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

**Wednesday 21 October 2020**

Afternoon (Time: 2 hours)

Paper Reference **9ST0/01**

**Statistics**

**Advanced**

**Paper 1: Data and Probability**

**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 9 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.

Turn over ►

P62709A

©2020 Pearson Education Ltd.

1/1/1/



Pearson

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

- 1 Shenandoah works at a medical practice. She wants to find out about patients' opinions on a planned extension to the practice and the disruption this may cause.

She plans to ask several questions to a sample of patients.

For **each** of the following methods, state **one** advantage, and **two** potential sources of bias.

- (a) Method A: Give a questionnaire, to be filled out whilst waiting, to each patient in the waiting room at the practice on a certain morning.

(i) one advantage

(1)

---

---

---

---

(ii) two potential sources of bias

(2)

---

---

---

---

---

---

---

---

---

---

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA





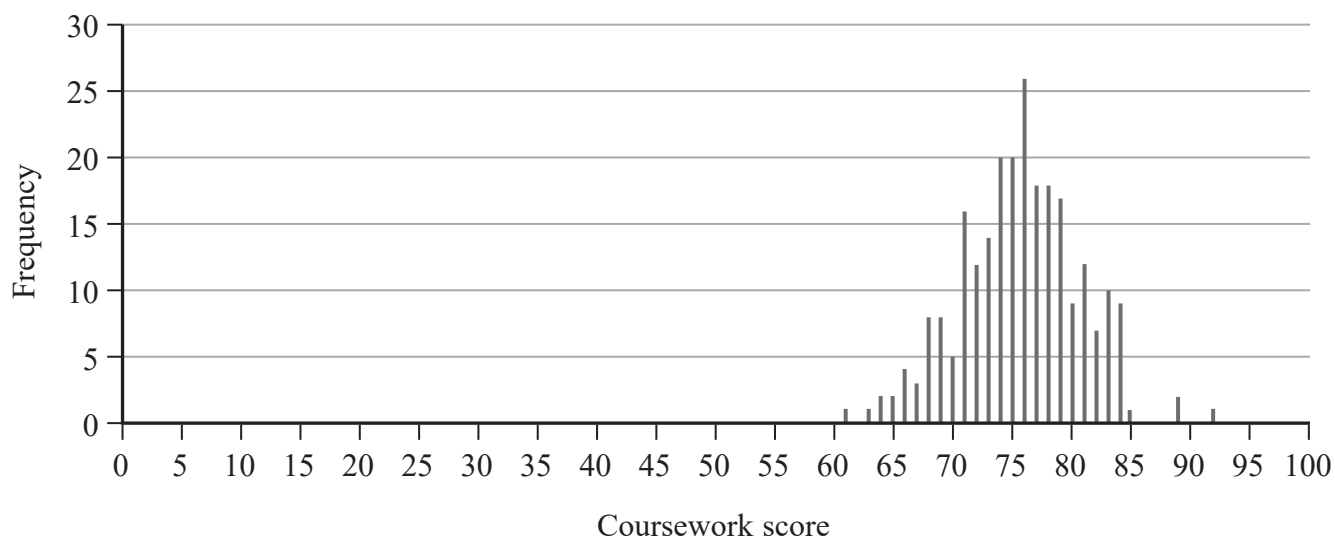




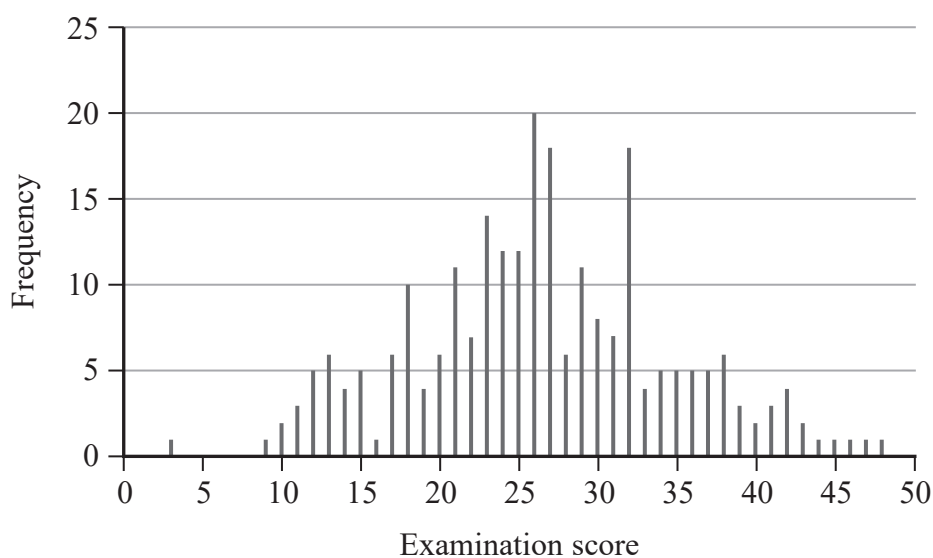
- 3 Sylvia runs a course at her local university. She decides to assess her students using both a coursework assignment and an examination.

Sylvia believes that the skills shown in the coursework are more important than those shown in the examination, so she decides to mark the coursework as an integer score out of 100, and the examination as an integer score out of 50

The scores for the first year of Sylvia's course assessments are given in **Figure 1** and **Figure 2**.



**Figure 1: Coursework scores in first year of assessments**



**Figure 2: Examination scores in first year of assessments**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**Question 3 continued**

Sylvia looks at **Figure 1** and **Figure 2** and decides that the normal distribution might be an appropriate model for each of the populations from which the data were taken. Tests for normality confirm her decision.

Hence, Sylvia uses normal distributions to model the scores.

(a) Describe:

(i) one advantage of Sylvia's model,

(1)

---

---

---

(ii) one disadvantage of Sylvia's model.

(1)

---

---

---



**Question 3 continued**

Summary statistics from the data are given below.

	<b>Coursework</b>	<b>Examination</b>
<b>Mean</b>	76.75	27.39
<b>Standard deviation</b>	4.76	8.19

- (b) Assuming that this year's students are a representative sample of all future students, use Sylvia's model to estimate the **coursework** score of a student at the 10th percentile of the coursework scores (the score below which 10% of the coursework scores lie).

You should give your answer to the nearest integer.

(3)

---

---

---

---

---

---

---

---

---

---

- (c) Assuming that this year's students are a representative sample of all future students, use Sylvia's model to estimate the **examination** score of a student at the 90th percentile of the examination scores (the score below which 90% of the examination scores lie).

You should give your answer to the nearest integer.

(1)

---

---

---

---

---

---

---

---

---

---





DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**Question 3 continued**

Each student's final score is the total of their coursework score and their examination score.

Student A scored at the 10th percentile in the coursework and at the 90th percentile in the examination.

Student B scored at the 90th percentile in the coursework with a score of 83, and at the 10th percentile in the examination with a score of 17. This gave Student B a total score of 100.

- (d) Do you think that Sylvia has successfully weighted the marks so that the total score reflects her goal that coursework scores are more important than the examination scores in the final score achieved?

Explain your answer.

(4)

---

---

---

---

---

---

---

---

---

---

- (e) What recommendations would you make for Sylvia when designing next year's assessments?

(2)

---

---

---

---

---

---

---

---

**(Total for Question 3 is 12 marks)**





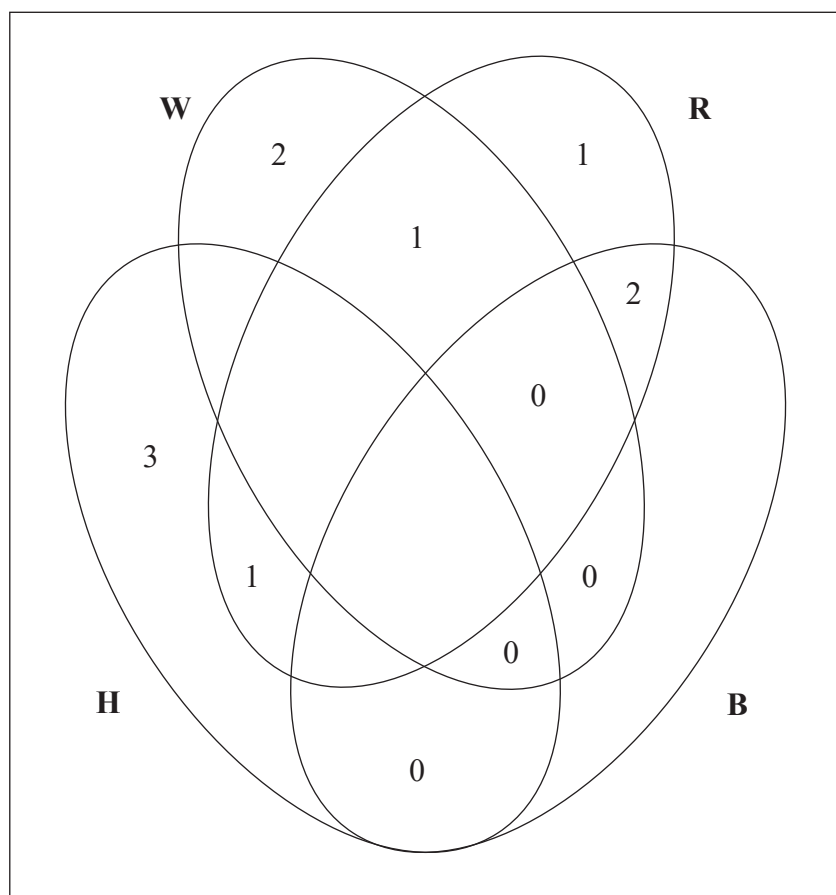


- 5 A game is played with a pack of 24 cards, each of which has a picture of a man or a woman on it, as well as his/her name.

If a card is picked at random, the following events are defined:

- H: The person in the picture is wearing a hat.  
 W: The person in the picture is a woman.  
 R: The person in the picture has red hair.  
 B: The person in the picture has a beard.

When the whole pack is considered, some of the frequencies of each event are presented in the Venn diagram in **Figure 4**.



**Figure 4**

In the pack of cards:

- there is one red-haired **man** with a hat and a beard,
- there are no red-haired women wearing a hat,
- there are nine people with a beard,
- exactly **half** of the women are wearing hats.

- (a) Fully complete the Venn diagram in **Figure 4**.

(5)



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**Question 5 continued**

A card is randomly drawn from the pack.

(b) Find the following probabilities:

(i)  $P(H \cap W' \cap R \cap B)$  (1)

---

---

---

(ii)  $P(H' \cap W' \cap R' \cap B')$  (1)

---

---

---

(iii)  $P(W \cap B)$  (1)

---

---

---

(iv)  $P(R|H)$  (2)

---

---

---

---

---

---



**Question 5 continued**

(c) Describe a person who satisfies  $H' \cap W' \cap R' \cap B'$

(1)

---

---

---

---

---

The game is played as follows:

- The cards are placed face-up on the table.
- Player Y selects a card at random without telling Player X what it is, and writes down the name of the person on the card on a piece of paper.
- Player X deduces which card Player Y selected by asking questions that require only yes/no answers.
- The object of the game is for Player X to identify the correct card with the fewest number of questions.

Statistically, the most efficient tactic for Player X is to ask a question that gets closest to dividing the cards in half with yes and no answers.

(d) Suggest a question based on **one** of the events H, W, R, or B, for Player X to ask to most effectively divide the cards.

(1)

---

---

---

---

---



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**Question 5 continued**

- (e) Suggest a new question that takes into account more than one of the events, which would be even more effective at dividing the cards than the question suggested in part (d).

(2)

---

---

---

---

---

---

---

---

**(Total for Question 5 is 14 marks)**







DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**Question 6 continued**

$T$  = total volume of concrete delivered (in  $\text{m}^3$ ) by  $k$  lorries.

$$T = X_1 + X_2 + \dots + X_k$$

(b) Specify fully the distribution of  $T$ .

(2)

---

---

---

---

---

---



**Question 6 continued**

The construction of the Three Gorges Dam in China required 27.2 million m<sup>3</sup> of concrete.

Li is in charge of planning for the construction of a similar-sized dam. She needs to estimate the total number,  $k$ , of lorry loads of concrete she must have delivered during construction in order to be **99% sure** she has enough to build the dam.

She makes the rough calculations shown in **Figure 5** on a pad of paper.

I need  $k$  lorry loads so I want:

$$P(T > 27\,200\,000) = \boxed{\phantom{000}}$$

So

$$z = \frac{27\,200\,000 - \boxed{\phantom{000}}}{\boxed{\phantom{000}}} = \boxed{\phantom{000}}$$

**Figure 5: Li's calculations and working (part 1)**

(c) Using your answer to (b), fill in the blanks in Li's working in **Figure 5**.

(3)

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---





- 7 Wilf is a town planner investigating the feasibility of building a new village in Oxfordshire. The village will not have its own shops, so he wants to model the potential shopping patterns of new residents.

He researches the surrounding towns and cities, and decides that the towns and cities shown in **Figure 7** are the most likely shopping destinations.

Name of town/city	Description	Size (population, 2011 census)	Distance from new village site (km)
Faringdon	Small town	7 121	5.1
Wantage	Small town	11 327	16.3
Witney	Medium town	27 522	17.5
Swindon	Large town	182 441	24.6
Oxford	City	150 200	25.7

**Figure 7**

He uses a version of the ‘gravity model of human interaction’ to estimate the proportion of the population’s shopping trips to a specific location (town/city).

His model uses the following formula:

$$P(T) = \frac{1}{1000} \times \frac{S}{D^2}$$

Where:

$T$  = Town/city

$S$  = Size of town/city (population)

$D$  = Distance to the town/city (km)

Wilf’s calculations have been started in **Figure 8**.

Town/city ( $T$ )	Size ( $S$ )	Distance ( $D$ )	Estimated proportion of shopping trips to the town/city ( $P(T)$ ) [to 4 decimal places]
<b>Faringdon</b>	7 121	5.1	0.2738
<b>Wantage</b>	11 327	16.3	0.0426
<b>Witney</b>	27 522	17.5	0.0899
<b>Swindon</b>	182 441	24.6	
<b>Oxford</b>	150 200	25.7	

**Figure 8**

- (a) Complete the table in **Figure 8**.

(2)



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**Question 7 continued**

- (b) Estimate the proportion of the population's shopping trips that will be spent in a small or medium town. (1)

---

---

---

By design, the estimated proportions in **Figure 8** add up to a total less than 1

- (c) Give a possible explanation for this. (1)

---

---

---

The proposed village has 300 houses planned.

The UK average number of people per household is 2.4

The UK average number of total household shopping trips per week is 1.9

- (d) Making any necessary assumptions, estimate how many more shoppers Faringdon should expect per week if the village is built.

You should clearly explain or show how you obtained your estimate. (3)

---

---

---

---

---

---

---

---

---

---













- 9 If a foul is committed in basketball, the player who had the foul committed against them is usually awarded **two** ‘free throws’. A point is scored from a ‘free throw’ if the player successfully throws the ball into the basket.

Aoife is a keen basketball player. Last year she achieved the statistics given in **Figure 9** when awarded two ‘free throws’.

		Second free throw	
		Scored	Did not score
First free throw	Scored	78	39
	Did not score	22	12

**Figure 9: Aoife’s free throw statistics last year**

- Professional basketball players should score from at least 70% of ‘free throws’.
- For professional basketball players, the outcome of their second ‘free throw’ is approximately independent of the outcome of their first ‘free throw’.

Determine whether the ‘free throw’ statistics in **Figure 9** suggest that Aoife is taking ‘free throws’ at a professional level.

You should support your argument with appropriate numerical evidence.

You should use your knowledge of probability theory in support of your answer.

You should **not** use a hypothesis test.

(9)





DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**BLANK PAGE**

Every effort has been made to contact copyright holders to obtain their permission for the use of copyright material. Pearson Education Ltd. will, if notified, be happy to rectify any errors or omissions and include any such rectifications in future editions.

