Please check the examination details below before entering your candidate information

Candidate surname _____ Other names _____

Centre Number ☐☐☐☐☐☐   Candidate Number ☐☐☐☐

Pearson Edexcel Level 3 GCE

Thursday 23 May 2024

Afternoon

Paper reference **8MA0/21**

Mathematics
Advanced Subsidiary
PAPER 21: Statistics

You must have:
Mathematical Formulae and Statistical Tables (Green), calculator

Total Marks _____

Candidates may use any calculator allowed by Pearson regulations. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or 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.
- Values from statistical tables should be quoted in full. If a calculator is used instead of tables the value should be given to an equivalent degree of accuracy.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- The total mark for this part of the examination is 30. There are 5 questions.
- 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.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.
- Check your answers if you have time at the end.

Turn over ▶

Pearson

1. A coach recorded the heights of some adult rugby players and found the following summary statistics.
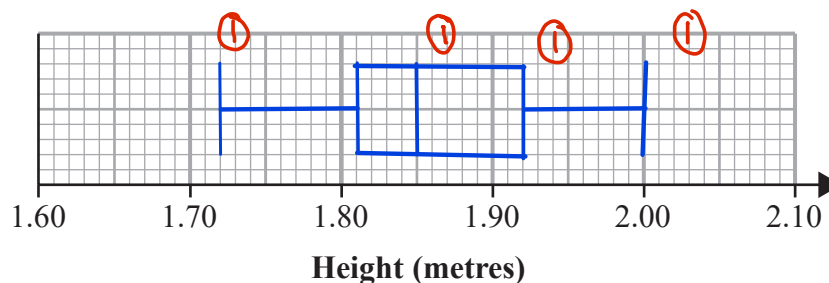
Median = 1.85 m

Range = 0.28 m

Interquartile range = 0.11 m

The coach also noticed that

- the height of the shortest player is 1.72 m

✱ • 25% of the players' heights are below the height of a player whose height is 1.81 m

Draw a box and whisker plot to represent this information on the grid below.

(4)

Height (metres)

from ✱, $Q_1 = 1.81$

$IQR = 0.11$
$IQR = Q_3 - Q_1$

$Q_3 - 1.81 = 0.11$
$Q_3 = 1.92$

Range = highest value - lowest value
$0.28 =$ highest value $- 1.72$
highest value $= 0.28 + 1.72 = 2.0$

Use the spare grid on page 3 if you need to redraw your box and whisker plot.

Question 1 continued

Only use this grid if you need to redraw your box and whisker plot.

Height (metres)
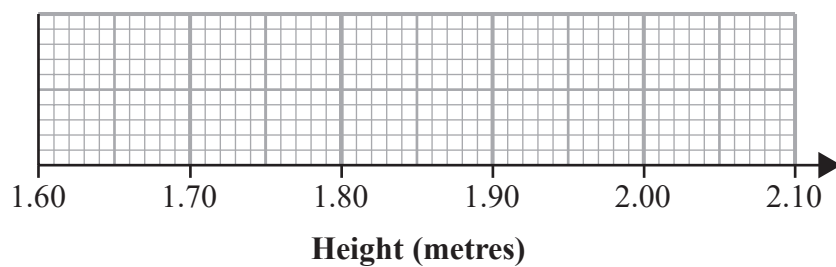
(Total for Question 1 is 4 marks)

P 7 5 6 8 0 A 0 3 1 6

2. Keith is studying the variable Daily Mean Wind Direction, in degrees, from the large data set.

   Keith summarised the data for Camborne from 1987 into 4 directions A, B, C and D representing North, South, East and West in some order.

Direction	A	B	C	D
Frequency	22	48	56	58

   (a) Using your knowledge of the large data set state, giving a reason, which direction A represents.

   (1)

   The entry for Hurn on 27th September 1987 was 999

   (b) State, giving a reason, what Keith should do with this value.

   (2)

a) A represents East, since the prevailing winds in Camborne are from the South, North or West. ①

b) Wind direction should be within the range 0 to 360, ① so the value should be ignored. ①

   Remember to study the large data set!

Question 2 continued

(Total for Question 2 is 3 marks)

3. Customers in a shop have to queue to pay.

The partially completed table below and partially completed histogram opposite, give information about the time, x minutes, spent in the queue by each of 112 customers one day.

Time in queue (x minutes)	Frequency
1–2	64
2–3	25
3–4	13
4–6	7
6–8	3

$112 - (64 + 25 + 13 + 3)$

①

No customer spent less than 1 minute or longer than 8 minutes in the queue.

(a) Complete the table.

(2)

(b) Complete the histogram.

(2)

Ting decides to model the **frequency density** for these 112 customers by a curve with equation

$$y = \frac{k}{x^2} \qquad 1 \leqslant x \leqslant 8$$
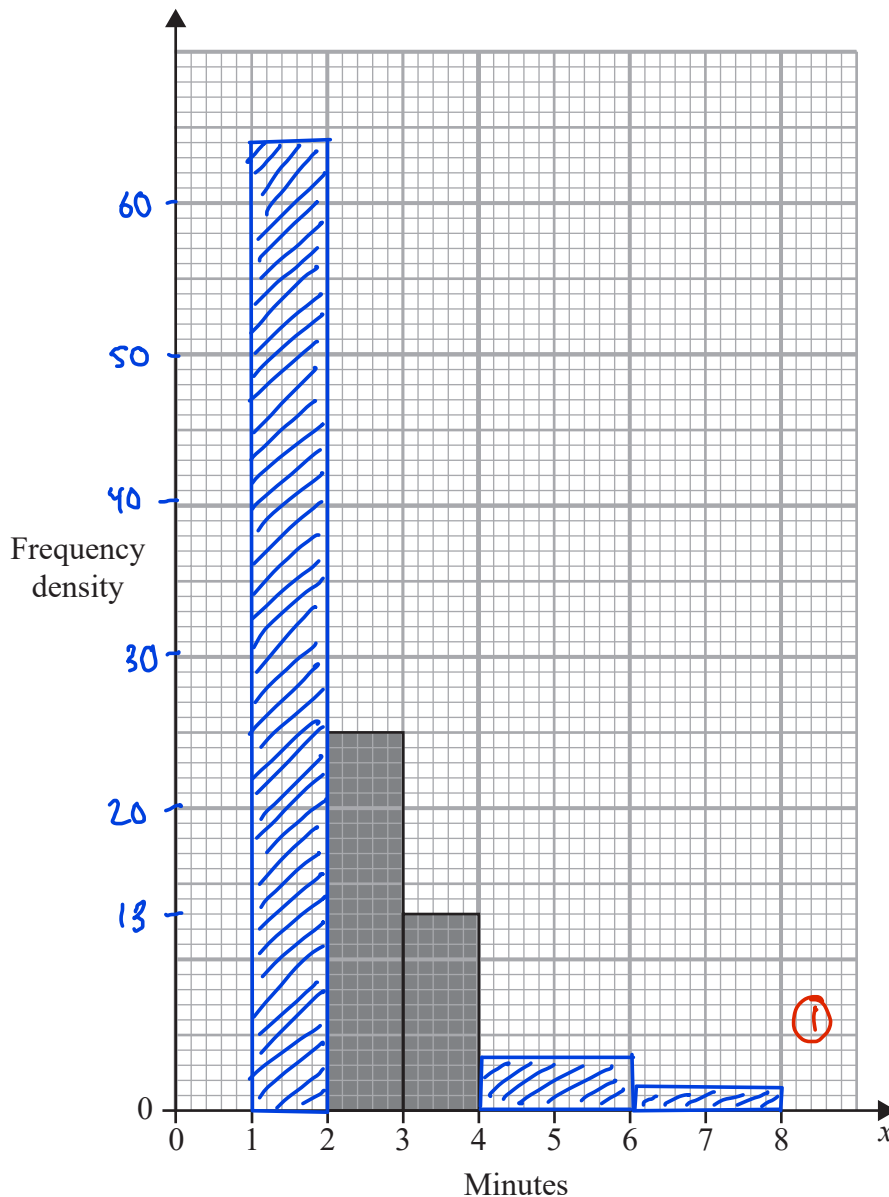
where k is a constant.

(c) Find the value of k

(3)

Question 3 continued

Frequency density (y-axis) vs Minutes (x-axis) histogram

a) 3 – 4 : frequency = 13 , class width (cw) = 1

∴ frequency density (fd) = $\frac{13}{1}$ = 13 ① can now label fd axis

2 – 3 frequency: 25 × 1 = 25

b) 1 – 2 : freq = 64, cw = 1 ⇒ fd = 64
4 – 6 : freq = 7, cw = 2 ⇒ fd = 7/2 = 3.5 ①
6 – 8 : freq = 3, cw = 2 ⇒ fd = 3/2 = 1.5

Turn over for a spare grid if you need to redraw your histogram.

P 7 5 6 8 0 A 0 7 1 6

Question 3 continued

c) fd: $y = \dfrac{k}{x^2}$ $1 \le x \le 8$
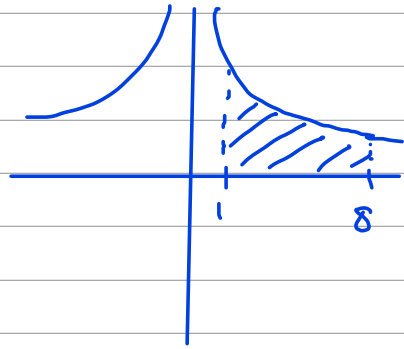
area of curve from $1 \le x \le 8$
= frequency

$\therefore \displaystyle\int_{1}^{8} kx^{-2}\, dx = 112$ ①

$\left[-kx^{-1} \right]_{1}^{8} = 112$

$\dfrac{-k}{8} + \dfrac{k}{1} = 112$ ①

$\dfrac{7}{8}k = 112$

$k = 128$ ①

Question 3 continued

Only use this grid if you need to redraw your histogram.

Frequency density

Minutes
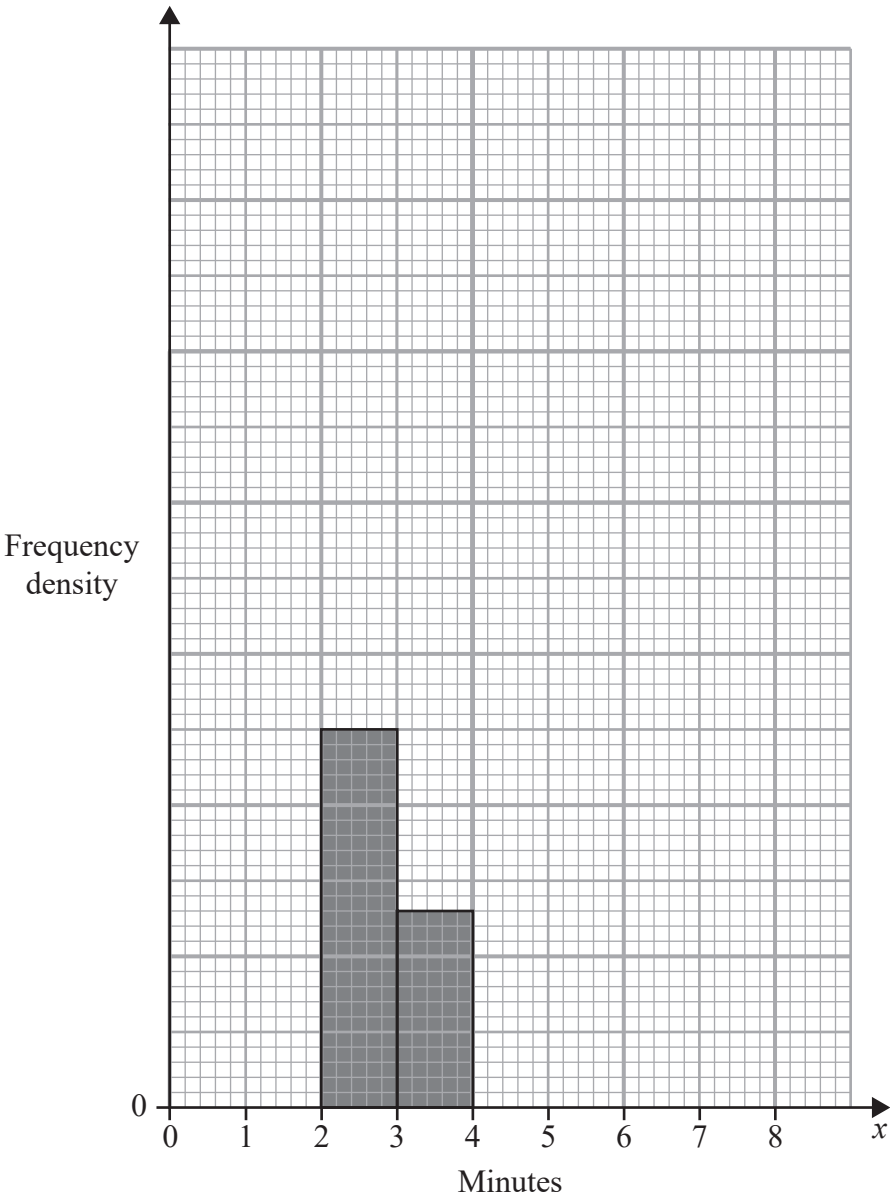
(Total for Question 3 is 7 marks)

Turn over

P 7 5 6 8 0 A 0 9 1 6

4. The random variable $X \sim B(27, 0.35)$

(a) Find   (i) $P(X = 10)$

    (ii) $P(12 \leqslant X < 15)$

(3)

Historical records show that the proportion of defective items produced by a machine is 0.12

Following a maintenance service of the machine, a random sample of 60 items is taken and 3 defective items are found.

1 tailed

(b) Carry out a suitable test to determine whether the proportion of defective items produced by the machine has decreased following the maintenance service. You should state your hypotheses clearly and use a 5% level of significance.

(4)

(c) Write down the *p*-value for your test in part (b)

(1)

a) (i) $P(X = 10) = 0.153588\ldots$
$$= 0.154 \ (3sf) \ \textcircled{1}$$

$\leftarrow X = 12, 13, 14$

(ii) $P(12 \leq X < 15) = P(12 \leq X \leq 14)$

$$= P(X \leq 14) - P(X \leq 11) \ \textcircled{1}$$

$$= 0.97707\ldots - 0.797603\ldots = 0.179 \ (3sf) \ \textcircled{1}$$

b) $H_0 : p = 0.12$ $H_1 : p < 0.12$ $\textcircled{1}$

Let D = the number of defective items in the sample.

Assume H_0 correct: $D \sim B(60, 0.12)$ $\textcircled{1}$

$P(D \leq 3) = 0.06013\ldots > 0.05$ $\textcircled{1}$

Insufficient evidence to reject H_0, so there is insufficient evidence that the proportion of defective items has decreased. $\textcircled{1}$

c) $0.0601 \ (3sf)$ $\textcircled{1}$

Question 4 continued

P75680A01116

Question 4 continued

Question 4 continued

(Total for Question 4 is 8 marks)

P75680A01316

5. A biased 4-sided spinner has the numbers 6, 7, 8 and 10 on it.

The discrete random variable X represents the score when the spinner is spun once and has the following probability distribution,

x	6	7	8	10
$P(X = x)$	0.5	0.2	q	q

where q is a probability.

(a) Find the value of q

(1)

Karen spins the spinner repeatedly until she **either** gets a 7 **or** she has taken 4 spins.

(b) Show that the probability that Karen stops after taking her 3rd spin is 0.128

(2)

The random variable S represents the number of spins Karen takes.

(c) Find the probability distribution for S

(4)

The random variable N represents the number of times Karen gets a 7

(d) Find $P(S > N)$

(1)

a) $0.5 + 0.2 + q + q = 1$

$2q = 0.3$

$q = 0.15$ ①

b) $P(\text{stops after 3rd spin}) = \text{not 7} \times \text{not 7} \times \text{yes 7}$ ①

$= 0.8 \times 0.8 \times 0.2$

$= 0.128$ ①

c) $S = 1, 2, 3, 4$ since game ends after 4 spins ①

$P(\text{stops after 1st spin}) = \text{yes 7}$

$= 0.2$

$P(\text{stops after 2nd spin}) = \text{not 7} \times \text{yes 7}$

$= 0.8 \times 0.2$

$= 0.16$ ①

Question 5 continued

$P(\text{stops after 4 spins}) = \text{not 7} \times \text{not 7} \times \text{not 7} \times \text{yes 7}$
$\underline{\text{OR}}$
$\rightarrow \text{not 7} \times \text{not 7} \times \text{not 7} \times \text{not 7}$

game ends after 4
spins even if Karen
doesnt spin a 7.

$= 0.8^3 \times 0.2 + 0.8^4$

$= 0.512$ ①

S	1	2	3	4
P(S=s)	0.2	0.16	0.128	0.512

①

d) $S = $ number of rolls
$N = $ number of sevens

for $S \le N$, $S = N = 1$

because there is always at least as many
rolls as there are sevens.

$\therefore P(S>N) = 1 - P(S \le N)$
$= 1 - P(S = N = 1)$
$= 1 - 0.2$
$= 0.8$ ①

Question 5 continued

(Total for Question 5 is 8 marks)

TOTAL FOR STATISTICS IS 30 MARKS