Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Paper reference **8MA0/21**

Mathematics

Advanced Subsidiary
PAPER 21: Statistics

October 2021 Shadow Set 1

You must have:
Mathematical Formulae and Statistical Tables, 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

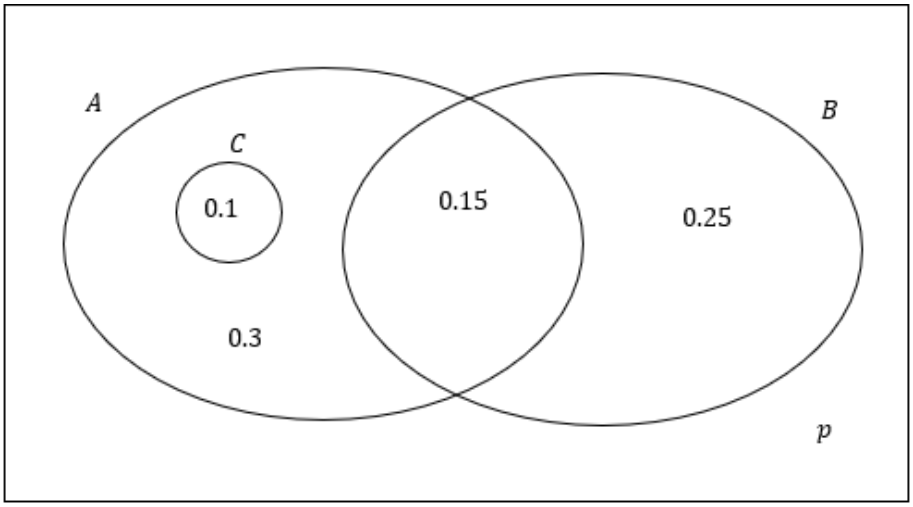
- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- The total mark for this part of the examination is 30. There are 5 questions.
- The marks for **each** question are shown in brackets
  – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.
- Good luck with your examination.

Turn over ▶

1.



The Venn diagram, where p is a probability, shows the 3 events A, B and C with their associated probabilities.

(a) Find the value of p.

(1)

(b) Write down a pair of mutually exclusive events from A, B and C.

(1)

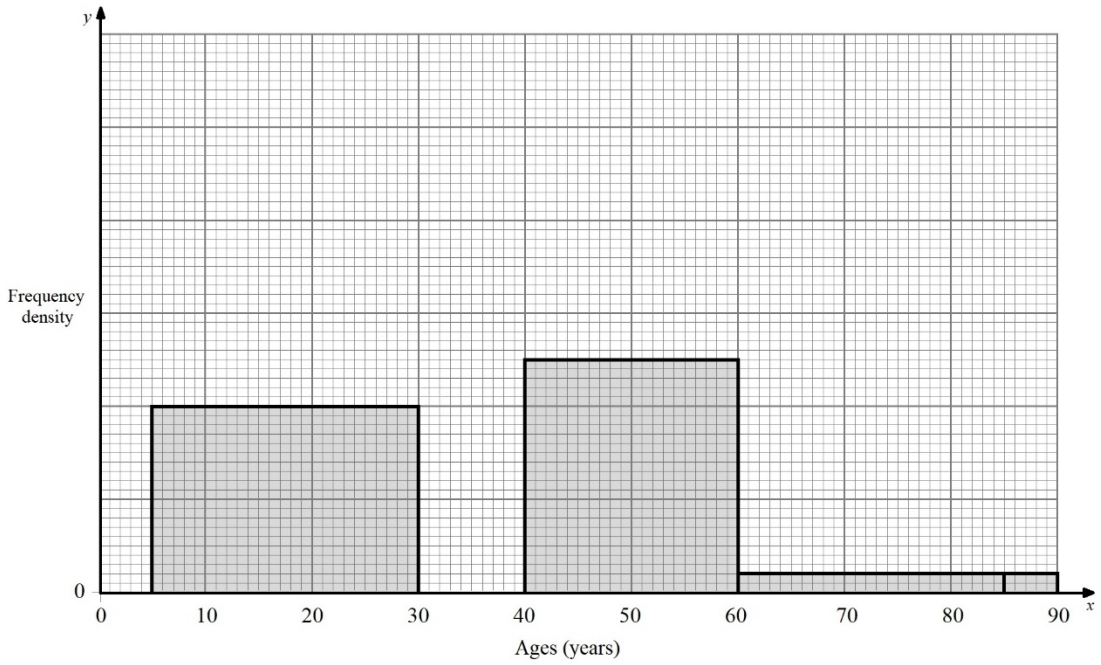
(Total for Question 1 is 2 marks)

---

2. The partially completed table and partially completed histogram give information about the ages of an audience in the theatre.

There were no audience members aged 90 or over.

Age (x years)	$0 \leqslant x < 5$	$5 \leqslant x < 30$	$30 \leqslant x < 40$	$40 \leqslant x < 60$	$60 \leqslant x < 85$	$85 \leqslant x < 90$
Frequency	5	50	55			1



(*a*) Complete the histogram.

(3)

(*b*) Use linear interpolation to estimate the median age.

(4)

An outlier is defined as a value greater than $Q_3 + 1.5 \times$ interquartile range.

Given that $Q_1 = 23.25$ and $Q_3 = 45.8$

(*c*) Determine, giving a reason, whether or not the oldest person in the audience could be considered as an outlier.

(2)

(Total for Question 2 is 9 marks)

---

3. Helen is studying one of the qualitative variables from the large data set for Cambourne from 1987.

She started with the data from 14th May and then took every 10th reading.

There were only 3 different outcomes with the following frequencies

Outcome	*A*	*B*	*C*
Frequency	12	5	1

(*a*) State the sampling technique Helen used.

(1)

(*b*) From your knowledge of the large data set

(i) suggest which variable was being studied,

(ii) state the name of outcome *B*.

(2)

George is also studying the same variable from the large data set for Cambourne from 2015. He started with the data from 1st May and then took every 10th reading and obtained the following

Outcome	*A*	*B*	*C*
Frequency	11	8	0

Helen and George decided they should examine all of the data for this variable for Cambourne from 2015 and obtained the following

Outcome	*A*	*B*	*C*
Frequency	119	58	6

(*c*) State what inference Helen and George could reliably make from their original samples about the outcomes of this variable at Cambourne, for the period covered by the large data set in 2015.

(1)

(Total for Question 3 is 4 marks)

---

4. An artist has a sack containing a large number of coloured tiles of which 12% are coloured green.

Duncan takes a random sample of 13 tiles from the sack to make a wall design.

(*a*) State a suitable binomial distribution to model the number of green tiles in Duncans design.

(1)

(*b*) Use this binomial distribution to find the probability that

(i) Duncan has 3 green tiles in the design,

(ii) there are at least 5 green tiles in the design.

(3)

(*c*) Comment on the suitability of a binomial distribution to model this situation.

(1)

After several people have used tiles from the sack, Duncan decides to test whether or not the proportion of green tiles has changed.
He takes a random sample of 60 tiles and finds 3 green tiles.

(*d*) Stating your hypotheses clearly, use a 5% significance level to carry out a suitable test for the artist.

(4)

(*e*) Find the *p*-value in this case.

(1)

(Total for Question 4 is 10 marks)

---

5. Two bags, **A** and **B**, each contain balls which are either red or yellow or green.

Bag **A** contains 5 red, 4 yellow and 2 green balls.
Bag **B** contains 5 red, n yellow and 4 green ball.

A ball is selected at random from bag **A** and placed into bag **B**.

A ball is then selected at random from bag **B** and placed into bag **A**.
The probability that bag **B** now contains an equal number of red, yellow and green balls is p.

Given that $p > 0$, find the possible values of n and p.

(5)

(Total for Question 5 is 5 marks)

TOTAL FOR STATISTICS IS 30 MARKS