Please check the examination details bel	ow before ente	ering your candidate information
Candidate surname		Other names
Centre Number Candidate Nu	ımber	
<b>Pearson Edexcel Level</b>	3 GCE	
Monday 18 October 2021 – Aftern	noon	
	Paper reference	9MA0/31
Mathematics		
Advanced		
PAPER 31: Statistics		
		J
You must have: Mathematical Formulae and Statistica	al Tahles (Gr	reen) calculator
Matricinatical Formulae and Statistica	110) 631001	Teerin, calculator

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 50. There are 6 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.

Turn over ▶







**1.** (a) State one disadvantage of using quota sampling compared with simple random sampling.

**(1)** 

In a university 8% of students are members of the university dance club.

A random sample of 36 students is taken from the university.

The random variable *X* represents the number of these students who are members of the dance club.

- (b) Using a suitable model for *X*, find
  - (i) P(X = 4)
  - (ii)  $P(X \ge 7)$

**(3)** 

Only 40% of the university dance club members can dance the tango.

(c) Find the probability that a student is a member of the university dance club and can dance the tango.

(1)

A random sample of 50 students is taken from the university.

(d) Find the probability that fewer than 3 of these students are members of the university dance club and can dance the tango.

**(2)** 





₫	
ARE	
S	
ČĽŇ	
	۲
č SA	
	۲
$\Diamond$ DU	ľ
800	Į
2	ļ
	į
8 🔯	
8 🔯	ľ
8 🔯	
8 🔯	
8 🔯	
WRIT	
WRIT	
WRIT	C R F F F F F F F
WRIT	
WRIT	
OT WRIT	
OT WRIT	
WRIT	
NOT WRIT	
NOT WRIT	
OT WRIT	
NOT WRIT	
NOT WRIT	
NOT WRIT	
NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	
O NOT WRIT	

Question 1 continued.	
	(Total for Question 1 is 7 marks)



- 2. Marc took a random sample of 16 students from a school and for each student recorded
  - the number of letters, x, in their last name
  - the number of letters, y, in their first name

His results are shown in the scatter diagram on the next page.

(a) Describe the correlation between x and y.

**(1)** 

Marc suggests that parents with long last names tend to give their children shorter first names.

(b) Using the scatter diagram comment on Marc's suggestion, giving a reason for your answer.

**(1)** 

The results from Marc's random sample of 16 observations are given in the table below.

x	3	6	8	7	5	3	11	3	4	5	4	9	7	10	6	6
у	7	7	4	4	6	8	5	5	8	4	7	4	5	5	6	3

(c) Use your calculator to find the product moment correlation coefficient between *x* and *y* for these data.

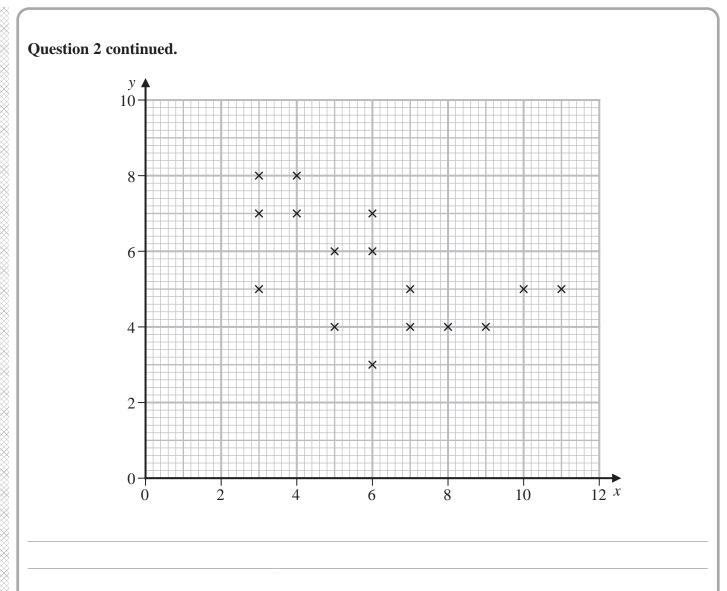
**(1)** 

(d) Test whether or not there is evidence of a negative correlation between the number of letters in the last name and the number of letters in the first name.

You should

- state your hypotheses clearly
- use a 5% level of significance

(3)





Question 2 continued.

>		×		×	⟨
^>	0		7	×	0
Κ	$\geq$	Š		L	à
/	и		Ą		5
>	C	7	ĕ	۹	ú
>	a	Ы	ú	И	
Κ	4	ь		i	ā
K	$\geq$		è	ä	à
/	И	Ľ		c	2
>	C	7	ē	Z	Ξ
>	G		ρ		7
Κ	7	8	и		è
K	$\geq$		×		
2	'n		ú	á	Ь
>	a	b			
`	Ğ		£	š	Z
Κ	ã	Ξ		Ξ	2
K	×	7		Т	3
>	G	ù		ù	ú
	G	Ē		Я	
۲	4	þ	×	í	p
K	у		×		
2	5	6	è	ê	š
>	C		ä	ρ	₹
>	1	ņ	я	ņ	2
X	a	į	à	ij	Ė
K			×		$\rangle$
2	ď		Ħ	r	'n
$\geq$	a	ь		ù	
ζ	8	le ke			2
<	1	k	of the second		
3		ŀ		2	3
3					
3					
3					
			THE REAL PROPERTY AND PERSONS ASSESSED.		
			THE REAL PROPERTY AND PERSONS ASSESSED.		
			THE REAL PROPERTY AND PERSONS ASSESSED.		
			N. Market Ball Co. of the		
			The party of the last of the l		
			THE RESERVE AND PARTY OF THE PA		
			THE RESERVE AND THE PARTY NAMED IN COLUMN TWO IS NOT THE PARTY NAMED IN COLUMN TWO IS		
			THE RESERVE AND THE PARTY OF TH		

Question 2 continued.	
(Total f	or Question 2 is 6 marks)
(10ttil 1	- Zueniu z in o inui ini)



3. Stav is studying the large data set for September 2015

He codes the variable Daily Mean Pressure, x, using the formula y = x - 1010

The data for all 30 days from Hurn are summarised by

$$\sum y = 214 \qquad \sum y^2 = 5912$$

(a) State the units of the variable x

**(1)** 

(b) Find the mean Daily Mean Pressure for these 30 days.

**(2)** 

(c) Find the standard deviation of Daily Mean Pressure for these 30 days.

**(3)** 

Stav knows that, in the UK, winds circulate

- in a **clockwise** direction around a region of **high** pressure
- in an anticlockwise direction around a region of low pressure

The table gives the Daily Mean Pressure for 3 locations from the large data set on 26/09/2015

Location	Heathrow	Hurn	Leuchars
Daily Mean Pressure	1029	1028	1028
<b>Cardinal Wind Direction</b>			

The Cardinal Wind Directions for these 3 locations on 26/09/2015 were, in random order,

W NE E

You may assume that these 3 locations were under a single region of pressure.

(d) Using your knowledge of the large data set, place each of these Cardinal Wind Directions in the correct location in the table. Give a reason for your answer.

**(2)** 

<<	
00000000	
- 888 <del>52</del> 68	
<u>va</u>	
$-\times$	
×ו••××	
XXXXXX	
$\times$	
$-\infty$	
	ı
- XX <del>AA</del> XX	
	ı
XXXXXX	П
	ı
	П
Ö	
ONC	
- XX XX XX I	
DO	ı
	l
	l
REA	
IS AREA	
IIS AREA	
IIS AREA	
IIS AREA	
THIS AREA	
THIS AREA	
THIS AREA	
N THIS AREA	
INTHIS AREA	
ITE IN THIS AREA	
ITE IN THIS AREA	
RITE IN THIS AREA	
RITE IN THIS AREA	
VRITE IN THIS AREA	
VRITE IN THIS AREA	
VRITE IN THIS AREA	
VRITE IN THIS AREA	
T WRITE IN THIS AREA	
T WRITE IN THIS AREA	
T WRITE IN THIS AREA	
T WRITE IN THIS AREA	
T WRITE IN THIS AREA	
NOT WRITE IN THIS AREA	
NOT WRITE IN THIS AREA	
O NOT WRITE IN THIS AREA	
O NOT WRITE IN THIS AREA	
O NOT WRITE IN THIS AREA	
O NOT WRITE IN THIS AREA	
NOT WRITE IN THIS AREA	

Question 3 continued.
(Total for Question 3 is 8 marks)



**4.** A large college produces three magazines.

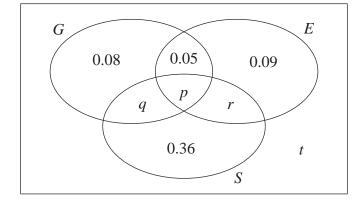
One magazine is about green issues, one is about equality and one is about sports. A student at the college is selected at random and the events G, E and S are defined as follows

G is the event that the student reads the magazine about green issues

*E* is the event that the student reads the magazine about equality

S is the event that the student reads the magazine about sports

The Venn diagram, where p, q, r and t are probabilities, gives the probability for each subset.



(a) Find the proportion of students in the college who read exactly one of these magazines.

**(1)** 

No students read all three magazines and P(G) = 0.25

- (b) Find
  - (i) the value of p
  - (ii) the value of q

**(3)** 

Given that  $P(S \mid E) = \frac{5}{12}$ 

- (c) find
  - (i) the value of r
  - (ii) the value of t

**(4)** 

(d) Determine whether or not the events  $(S \cap E')$  and G are independent. Show your working clearly.

**(3)** 



<b>****</b>	
XXXXX	
	`
>>>>> l	
XXXX	
(a)	-
	-
$\times$ v $\alpha$ $\times$	_
XX225XX	
	l _
$\times \times \times$	
	-
	-
$\otimes \odot \otimes$	-
Ž	
00	-
$\otimes \triangle \otimes \Box$	
***** I	l _
XXXX	
***** I	
>>>>	-
XXXX	
XXXX	
XXXXX	-
***** I	
>>>> I	
XXXX	-
>>>>	
XXXX	
XXXX	-
	_
$\times$	
4	
×w×	l _
<b>***</b>	
$\bowtie \bowtie \bowtie$	-
	-
	-
$\times \times \times \times$	
$\otimes$	-
⊗©×	-
2	
ONO	_
ON OC	-
DO NO	-   -   _
DO NO	-
XXX	-
$\times\!\!\times\!\!\times\!\!\times$	-   -   -
	-
	-
	-
	-
	-
	-
S AREA	-
S AREA	-
HIS AREA	-
HIS AREA	-
N THIS AREA	
N THIS AREA	
E IN THIS AREA	
E IN THIS AREA	
E IN THIS AREA	
RITE IN THIS AREA	
WRITE IN THIS AREA	
T WRITE IN THIS AREA	
T WRITE IN THIS AREA	
IOT WRITE IN THIS AREA	
NOT WRITE IN THIS AREA	
O NOT WRITE IN THIS AREA	
IOT WRITE IN THIS AREA	

Question 4 continued.	



Question 4 continued.

	>		>		$\geq$
	>				7
>	۲,	à	Ę		e
ι	A		2		2
0	S		3	7	9
>	a	ŧ	q	ρ	1
١.	Α	b		ь	а
	S	Z	2	3	Ζ
>	q	r	7	ρ	₹
	d	h	μ	×	ø
	>		2	4	à
>	Ú	į	5		₹
	2	7	9	۹	è
	>				Z
>	ς	2	ς,	à	Ĺ
5	4	ĸ.	А	ĸ	3
		₹	7		5
2	ű	μ	۰	ρ	80,
/	Ó	š	4	š	ú
	>	r	1		3
?	Ì	ú		ы	ĺ
	á	í			
C	A		è	á	ù
7	V	F	١,	1	Ι,
					C
	Þ	٥	۰		ø
	١	á		۳	۲,
	0	9	e	,	μ
	a	ĥ	×	'n	ni
/	١,		/	/	V
	۲,	۵	۷.		4
	K	I.	j	L	ã
	ß	۲	7	,	۹,
	a		∠	2	≤
	4		7	7	7
	S		ì	_	_
	œ	7	₹	7	₹
ς	4		×		ρ
	И	Ľ	4	Е	4
	G	2		7	↸
١	7	7	7	9	Νí
	K			c	5
$\geq$	C	2	S		Ń
S	Λ	۹	7	5	2
	À	6			×
	q		_	2	'n
\	4		7	2	7
	>	ú	ø	ø	×
2	И	ŕ		2	S)
>	0	9	6	à	g
	à	í	3	S	4
2	ς	7	S		ĸ,
>	ű			š	6
	2	₹	7	₹	7
7	5		À	Ζ	S
>	G	ø	₹	9	ĸ
	Α	ĸ.	2	S	a
		7	9	ø	۲,
>		à	2		6
`	4	ď.			坡
	y	Ħ	ņ	ø	峴
>					
	2				7
	>				$\geq$
>	C				<.
1	2	5	2		
	>				×

Question 4 continued.	
(Total for Question 4 is 11 marks)	



- 5. The heights of females from a country are normally distributed with
  - a mean of 166.5 cm
  - a standard deviation of 6.1 cm

Given that 1% of females from this country are shorter than  $k \, \text{cm}$ ,

(a) find the value of k

**(2)** 

(b) Find the proportion of females from this country with heights between 150 cm and 175 cm

**(1)** 

A female, from this country, is chosen at random from those with heights between 150 cm and 175 cm

(c) Find the probability that her height is more than 160 cm

**(4)** 

The heights of females from a different country are normally distributed with a standard deviation of 7.4 cm

Mia believes that the mean height of females from this country is less than 166.5 cm

Mia takes a random sample of 50 females from this country and finds the mean of her sample is 164.6 cm

(d) Carry out a suitable test to assess Mia's belief.

You should

- state your hypotheses clearly
- use a 5% level of significance

**(4)** 

ITE IN THIS AREA
S
五
WRI
E
DO NOT
<u></u>
ă
111
00
V2
IN THIS AREA
$\times \times \times$
imesoc $ imes$
WR
OT WR
NOT WR
O NOT WR
DO NOT WRITE!
DO NOT WR
IN THIS AREA
TE IN THIS AREA
TE IN THIS AREA
TE IN THIS AREA
TE IN THIS AREA
TE IN THIS AREA
NOT WRITE IN THIS AREA
TE IN THIS AREA

Question 5 continued.	



Question 5 continued.

Question 5 continued.
(Total for Orestion 5 in 11 months)
(Total for Question 5 is 11 marks)



**6.** The discrete random variable X has the following probability distribution

x	а	b	c
P(X = x)	$\log_{36} a$	$\log_{36} b$	$\log_{36} c$

where

- a, b and c are distinct integers (a < b < c)
- all the probabilities are greater than zero
- (a) Find
  - (i) the value of a
  - (ii) the value of b
  - (iii) the value of c

Show your working clearly.

**(5)** 

The independent random variables  $\boldsymbol{X}_{1}$  and  $\boldsymbol{X}_{2}$  each have the same distribution as  $\boldsymbol{X}$ 

(b) Find  $P(X_1 = X_2)$ 

**(2)** 



$\sim$	
$\approx$	XX
22	XX.
X	XX
$\sim$	$\times\!\!\!\times\!\!\!\!\circ$
XX	889
××	$\times \times$
XX	$\times\!\!\times\!\!\times$
XX	$\propto \times$
X	$\sim$
$\Rightarrow \Rightarrow$	$\otimes$
$\otimes$	<b>3</b>
X	NIA.
$\times$	
X	-
X	
××	$\times$
××	$\times$
X.	VA:
×	
×	-
×X	MM
831	$\times$
XX	V.
××	$\times \times$
XX	
XX	
XX.	$\propto$
X	ĽĽĽK
$\approx$	
$\approx$	-
$\times$	
X	2.7
$\sim$	
$\sim$	
××	$\times$
$\propto$	$\sim$
	$\simeq$
88	
SS	DO NOT WRITE IN THIS AREA
58	
X	
$\otimes$	$\sim$
$\otimes$	الان
$\times$	
X	
X	XX
$\langle \rangle$	XX:
$\propto$	XX
$\langle \rangle$	XX
$\times \times$	$\times\!\!\!\times\!\!\!\!\sim$
$\langle \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \!$	$\times\!\!\!\times\!\!\!\circ$
X	$\times\!\!\times$
$\times\!\!\times$	$\times\!\!\times\!\!\!\sim$
××	$\times \times$
×>	$\sim$
X	$\propto \times$
>>>	$\times \times$
$\sim$	$\diamond\diamond$
>>	$\otimes \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \! \!$
>>	$\otimes \!\! \times$
$\times$	$\propto \times$
X	
X	XX
~~	XX.
$\times\!\!\times$	$\times\!\!\times\!\!\!\circ$
$\times\!\!\times$	$\times\!\!\!\times\!\!\!\!\circ$
××	$\times \times$
××	$\times \times$
	$\times \times$
585	
$\otimes$	
$\overset{\times}{\otimes}$	4
$\stackrel{\times}{\otimes}$	A H
	N. A.
	AREA
	S. A. R. H.A.
	SAREA
	IS AREA
	LIS AREA
	HIS AREA
	THIS AREA
	THIS AREA
	VIEWS AREA
	NATES AREA
	INTERS AREA
	TINATE STATES
	THE VIEW
	TENNIH SAREA
	STEIN SHIS AREA
	RIE INTHIS AREA
	VENTERN YEARS AREA
	WRITE IN YELS AREA
	WRITE IN THIS AREA
	TWRITE IN THIS AREA
	NEW STREET, NEW ST
	OT WRITE IN THIS AREA
	OT WATER WITH SAREA
	NOT WRITE INTERS AREA
	NOT WRITE IN THIS AREA
	SNOT WRITEIN YHIS AREA
	ONOT WRITE IN THIS AREA
	SONOT WRITE AN ALMS AREA
	DONOTWRITE INTERS AREA
	DO NOT WRITE IN THIS AREA
	DO NOT WRITE IN THIS AREA
	OONOT WRITEIN THIS AREA
	OONOT WRITE'N THIS AREA
	OONOT WRITE IN THIS AREA
	NOT WRITE IN THIS AREA
	SONOT WRITE IN THIS AREA
	DO NOT WRITE IN THIS AREA
	DO NOT WRITE TO THIS AREA
	SONOT WRITEIN THIS AREA
	DO NOT WRITE IN THIS AREA
	DO NOT WRITE IN THIS AREA
	OONOT WRITEHN THIS AREA
	SONOT WRITE IN THIS AREA
	SONOT WRITE IN THIS AREA
	DO NOT WRITE'IN THIS AREA
	O NOT WRITE IN THIS AREA
	SONOT WRITE IN THIS AREA
	DO NOT WRITE IN THIS AREA
	DONOT WRITE IN THIS AREA
	DO NOT WRITE IN THIS AREA
	SO NOT WRITE'N THIS AREA
	DONOT WRITE'N THIS AREA
	<b>4</b>
	<b>4</b>
	<b>4</b>
	REA
	REA
	AREA
	AREA
	AREA
	REA

				ĺ	
				3	
				į	
				ì	
			2		
	7	9	è	ζ	
	Š	Ì	ì	2	
	Š	Ė	N IN		
2			100		
Š		200			
Š					
Ş					
Š					
Š		ц	ø	щ	
Š		ц	ø	щ	
3		ц	ø	щ	
3		ц	ø	щ	
3		ц		щ	
		ц	ø	щ	
		ц	ø	щ	
		ц	ø	щ	
		ц	ø	щ	
		ц	ø	щ	
		ц	ø	щ	
		ц	ø	щ	
			ø		

Question 6 continued.	



Question 6 continued.
(Total for Question 6 is 7 marks)
TOTAL FOR STATISTICS IS 50 MARKS

