

Please check the examination details below before entering your candidate information

Candidate surname

Other names

**Pearson Edexcel  
Level 3 GCE**

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

**Monday 19 October 2020**

Afternoon

Paper Reference **9MA0/31**

**Mathematics**

**Advanced**

**Paper 31: 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 50. 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.

Turn over ►

P66788A

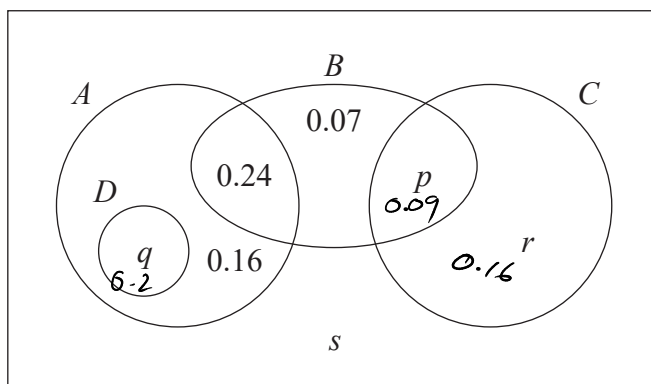
©2020 Pearson Education Ltd.

1/1/1/1/1/



Pearson

1. The Venn diagram shows the probabilities associated with four events,  $A$ ,  $B$ ,  $C$  and  $D$



- (a) Write down any pair of mutually exclusive events from  $A$ ,  $B$ ,  $C$  and  $D$  (1)

Given that  $P(B) = 0.4$

- (b) find the value of  $p$  (1)

Given also that  $A$  and  $B$  are independent

- (c) find the value of  $q$  (2)

Given further that  $P(B'|C) = 0.64$

- (d) find (4)
- (i) the value of  $r$
- (ii) the value of  $s$

a)  $A$  and  $C$  (or  $D$  and  $B$  or  $D$  and  $C$ )

b)  $P(B) = 0.4$   
 $0.4 = 0.24 + 0.07 + p$   
 $p = 0.09$

c)  $P(A) \times P(B) = P(A \cap B)$

$P(A) \times 0.4 = 0.24$

$P(A) = 0.6$

$q = 0.6 - 0.24 - 0.16$   
 $= 0.2$



Question 1 continued

$$P(B'|C) = 0.64$$

$$P(B'|C) = \frac{P(B' \cap C)}{P(C)}$$

$$0.64 = \frac{r}{p+r}$$

$$0.64 = \frac{r}{0.09+r}$$

$$0.64(0.09+r) = r$$

$$0.0576 + 0.64r = r$$

$$0.0576 = 0.36r$$

$$r = \frac{0.0576}{0.36}$$

$$= \underline{\underline{0.16}}$$

$$\text{ii/ } S = 1 - 0.16 - 0.24 - 0.07 - 0.09 - 0.16 - 0.2$$

$$= \underline{\underline{0.08}}$$

(Total for Question 1 is 8 marks)





Question 2 continued

a/ Negative Correlation

b/ Rainfall (mm)

$$c/ \begin{aligned} H_0: & \rho = 0 \\ H_1: & \rho \neq 0 \end{aligned}$$

$$n=30 \quad \text{critical value} = -0.3610$$

$$-0.377 < -0.3610$$

There is evidence to suggest there is a correlation between sunshine and humidity

d/ Humidity is high and there is a negative correlation  
I would expect a lower than average amount of sunshine



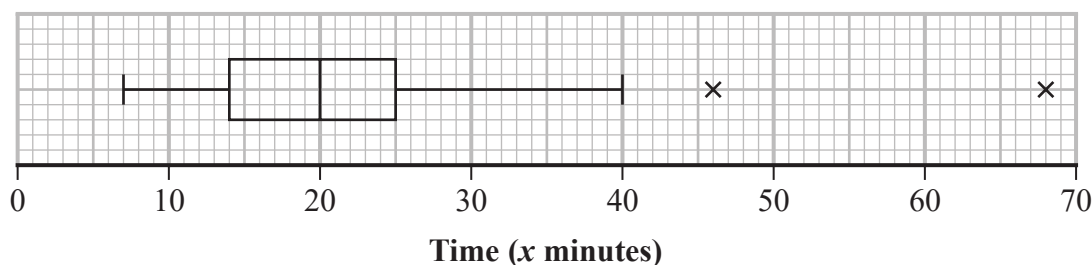




3. Each member of a group of 27 people was timed when completing a puzzle.

The time taken,  $x$  minutes, for each member of the group was recorded.

These times are summarised in the following box and whisker plot.



- (a) Find the range of the times. (1)

- (b) Find the interquartile range of the times. (1)

For these 27 people  $\sum x = 607.5$  and  $\sum x^2 = 17\,623.25$

- (c) calculate the mean time taken to complete the puzzle, (1)

- (d) calculate the standard deviation of the times taken to complete the puzzle. (2)

Taruni defines an outlier as a value more than 3 standard deviations above the mean.

- (e) State how many outliers Taruni would say there are in these data, giving a reason for your answer. (1)

Adam and Beth also completed the puzzle in  $a$  minutes and  $b$  minutes respectively, where  $a > b$ .

When their times are included with the data of the other 27 people

- the median time increases
  - the mean time does not change
- (f) Suggest a possible value for  $a$  and a possible value for  $b$ , explaining how your values satisfy the above conditions. (3)
- (g) Without carrying out any further calculations, explain why the standard deviation of all 29 times will be lower than your answer to part (d). (1)

$$a) \quad 68 - 7 = \underline{\underline{61}}$$

$$b) \quad 25 - 14 = \underline{\underline{11}}$$





Question 3 continued

$$c/ \frac{\sum x}{n} = \frac{607.5}{27} = \underline{\underline{22.5}}$$

$$d/ \sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$\sigma = \sqrt{\frac{17623.25}{27} - (22.5)^2}$$

$$\sigma = \underline{\underline{12.1}}$$

$$e/ 22.5 + 3(12.1) = 58.8$$

One outlier

f/ Mean stays the same  $\therefore$  mean of times must be 22.5  
and both above median (20)

$\therefore$  22 and 23  
a = 23    b = 22

g/ Both values will be close to the mean (both within 1 standard deviation of the mean)  $\therefore$   $\sigma$  will decrease







4. The discrete random variable  $D$  has the following probability distribution

$d$	10	20	30	40	50
$P(D = d)$	$\frac{k}{10}$	$\frac{k}{20}$	$\frac{k}{30}$	$\frac{k}{40}$	$\frac{k}{50}$

where  $k$  is a constant.

- (a) Show that the value of  $k$  is  $\frac{600}{137}$  (2)

The random variables  $D_1$  and  $D_2$  are independent and each have the same distribution as  $D$ .

- (b) Find  $P(D_1 + D_2 = 80)$   
Give your answer to 3 significant figures. (3)

A single observation of  $D$  is made.

The value obtained,  $d$ , is the common difference of an arithmetic sequence.

The first 4 terms of this arithmetic sequence are the angles, measured in degrees, of quadrilateral  $Q$

- (c) Find the exact probability that the smallest angle of  $Q$  is more than  $50^\circ$  (5)

$$a) \quad \frac{k}{10} + \frac{k}{20} + \frac{k}{30} + \frac{k}{40} + \frac{k}{50} = 1$$

$$\frac{137}{600} k = 1$$

$$k = \frac{600}{137}$$

$$b) \quad P(30 \text{ and } 50) = \frac{1}{30} \times \frac{600}{137} \times \frac{1}{50} \times \frac{600}{137} = 0.0128$$

$$P(50 \text{ and } 30) = 0.0128$$

$$P(40 \text{ and } 40) = \frac{1}{40} \times \frac{600}{137} \times \frac{1}{40} \times \frac{600}{137} = 0.0120$$

$$2(0.0128) + 0.0120 = \underline{\underline{0.0376}}$$

$$c) \quad S_n = \frac{n}{2} (2a + (n-1)d)$$

$$360 = \frac{4}{2} (2a + 3d)$$



Question 4 continued

$$360 = 2(2a + 3d)$$

$$180 = 2a + 3d$$

if  $d = 10$   $180 = 2a + 3(10)$

$$150 = 2a$$

$$a = 75^\circ$$

$d = 20$   $180 = 2a + 3(20)$

$$120 = 2a$$

$$a = 60^\circ$$

$d = 30$   $180 = 2a + 3(30)$

$$90 = 2a$$

$$a = 45^\circ$$

$d = 10$  or  $d = 20$

$$\frac{k}{10} + \frac{k}{20} = \underline{\underline{\frac{90}{137}}}$$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA







5. A health centre claims that the time a doctor spends with a patient can be modelled by a normal distribution with a mean of 10 minutes and a standard deviation of 4 minutes.

- (a) Using this model, find the probability that the time spent with a randomly selected patient is more than 15 minutes.

(1)

Some patients complain that the mean time the doctor spends with a patient is more than 10 minutes.

The receptionist takes a random sample of 20 patients and finds that the mean time the doctor spends with a patient is 11.5 minutes.

- (b) Stating your hypotheses clearly and using a 5% significance level, test whether or not there is evidence to support the patients' complaint.

(4)

The health centre also claims that the time a dentist spends with a patient during a routine appointment,  $T$  minutes, can be modelled by the normal distribution where  $T \sim N(5, 3.5^2)$

- (c) Using this model,

- (i) find the probability that a routine appointment with the dentist takes less than 2 minutes

(1)

- (ii) find  $P(T < 2 \mid T > 0)$

(3)

- (iii) hence explain why this normal distribution may not be a good model for  $T$ .

(1)

The dentist believes that she cannot complete a routine appointment in less than 2 minutes.

She suggests that the health centre should use a refined model only including values of  $T > 2$

- (d) Find the median time for a routine appointment using this new model, giving your answer correct to one decimal place.

(5)

$$\mu = 10 \quad \sigma = 4$$

$$\text{Normal CD} \quad P(X > 15) = \underline{\underline{0.106}}$$

$$\begin{aligned} \text{b/ } H_0: \mu &= 10 & \mu &= 10 & \sigma &= \frac{4}{\sqrt{20}} \\ H_1: \mu &> 10 & & & & \end{aligned}$$

$$P(X > 11.5) = 0.0468$$

$0.0468 < 0.05$  There is evidence to suggest the





Question 5 continued

mean time, is greater than 10 min  
the doctor spends with each patient

c/  $\mu = 5 \quad \sigma = 3.5$

$$P(X < 2) = \underline{\underline{0.196}}$$

ii/  $P(X > 0) = 0.923$

$$P(0 < X < 2) = 0.119$$

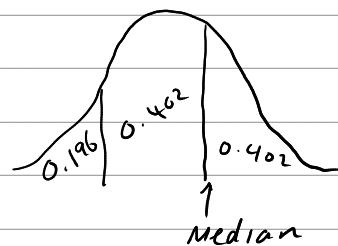
$$P(X < 2 | X > 0) = \frac{P(0 < X < 2)}{P(X > 0)} = \frac{0.119}{0.923} \\ = \underline{\underline{0.129}}$$

iii/ There will be patients who saw the dentist for a negative amount of time.

$P(X < 2)$  should equal  $P(X < 2 | X > 0)$  but the answers are not close.

d/  $P(X < 2) = 0.196$

$$P(X > 2) = 0.804$$



$$0.402 + 0.196 = 0.598$$

Inverse normal Area = 0.598

$$X = \underline{\underline{5.87}}$$







