

**Paper Reference 9MA0/31**  
**Pearson Edexcel**  
**Level 3 GCE**

**Mathematics**  
**Advanced**  
**Paper 31: Statistics**

**Monday 19 October 2020 – Afternoon**

**MATERIALS REQUIRED FOR  
EXAMINATION**

**Mathematical Formulae and Statistical  
Tables (Green), calculator**

**ITEMS INCLUDED WITH QUESTION  
PAPER**

**Answer Book**  
**Diagram Book**

**Y66788A**

**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**

**In the boxes on the Answer Book and on the Diagram Book, write 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 Answer Book or on the separate diagrams – there may be more space than you need.**

**Do NOT write on the Question Paper.**

**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.**

**Turn over**

## **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.**

1. Refer to the diagram for Question 1 in the Diagram Book.

It is a Venn diagram which 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 mark)

Given that  $P(B) = 0.4$

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

(continued on the next page)

Turn over

**1. continued.**

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

**(c) find the value of  $q$   
(2 marks)**

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

**(d) find**

**(i) the value of  $r$**

**(ii) the value of  $s$**

**(4 marks)**

**(Total for Question 1 is 8 marks)**

---

**Turn over**

- 2. Refer to the diagram for Question 2 in the Diagram Book.**

**A random sample of 15 days is taken from the large data set for Perth in June and July 1987**

**The scatter diagram in the Diagram Book displays the values of two of the variables for these 15 days.**

- (a) Describe the correlation.  
(1 mark)**

**(continued on the next page)**

**Turn over**

**2. continued.**

**The variable on the X-axis is**

**Daily Mean Temperature measured  
in °C**

**(b) Using your knowledge of the  
large data set,**

**(i) suggest which variable is on  
the y-axis,**

**(ii) state the units that are used  
in the large data set for this  
variable.**

**(2 marks)**

**(continued on the next page)**

**Turn over**



**2. continued.**

**Stav believes that there is a correlation between Daily Total Sunshine and Daily Maximum Relative Humidity at Heathrow.**

**He calculates the product moment correlation coefficient between these two variables for a random sample of 30 days and obtains  $r = -0.377$**

**(continued on the next page)**

**Turn over**

**2. continued.**

**(c) Carry out a suitable test to investigate Stav's belief at a 5% level of significance.**

**State clearly**

- **your hypotheses**
- **your critical value**

**(3 marks)**

**(continued on the next page)**

**Turn over**

**2. continued.**

**On a random day at Heathrow the  
Daily Maximum Relative Humidity  
was 97%**

- (d) Comment on the number of  
hours of sunshine you would  
expect on that day, giving a  
reason for your answer.  
(1 mark)**

**(Total for Question 2 is 7 marks)**

---

**Turn over**

- 3. Refer to the diagram for Question 3 in the Diagram Book.**

**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 box and whisker plot in the Diagram Book.**

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

**(continued on the next page)**

**Turn over**

**3. continued.**

**(b) Find the interquartile range of the times.**

**(1 mark)**

**For these 27 people**

$$\sum x = 607 \cdot 5 \text{ and}$$

$$\sum x^2 = 17\,623 \cdot 25$$

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

**(1 mark)**

**(continued on the next page)**

**Turn over**

**3. continued.**

**Remember:**

**For these 27 people**

$$\sum x = 607 \cdot 5 \text{ and}$$

$$\sum x^2 = 17\,623 \cdot 25$$

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

**(continued on the next page)**

**Turn over**

**3. continued.**

**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 mark)**

**(continued on the next page)**

**3. continued.**

**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 marks)**

**(continued on the next page)**

**Turn over**



**3. continued.**

- (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 mark)**

**(Total for Question 3 is 10 marks)**

---

**Turn over**

4. Refer to the table for Question 4 in the Diagram Book.

The discrete random variable **D** has the probability distribution shown in the Diagram Book where **k** is a constant.

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

(continued on the next page)

**4. continued.**

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

**(b) Find**

$$\mathbf{P(D_1 + D_2 = 80)}$$

**Give your answer to  
3 significant figures.  
(3 marks)**

**(continued on the next page)**

**Turn over**

**4. continued.**

**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 marks)**

**(Total for Question 4 is 10 marks)**

---

**Turn over**

- 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 mark)**

**(continued on the next page)**

**Turn over**

**5. continued.**

**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 marks)**

**(continued on the next page)**

**Turn over**

**5. continued.**

**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**

$$\mathbf{T \sim N(5, 3 \cdot 5^2)}$$

**(c) Using this model,**

**(i) find the probability that a routine appointment with the dentist takes less than**

**2 minutes**

**(1 mark)**

**(continued on the next page)**

**Turn over**

**5. (c) continued.**

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

**(3 marks)**

**(iii) hence explain why this  
normal distribution may not  
be a good model for  $T$   
(1 mark)**

**(continued on the next page)**



**5. continued.**

**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 marks)**

**(Total for Question 5 is 15 marks)**

---

**Turn over**

---

**TOTAL FOR STATISTICS IS 50 MARKS**  
**END OF PAPER**

---