

**Paper Reference(s) 9MA0/31**

**Pearson Edexcel Level 3 GCE**

**Mathematics**

**Advanced**

**PAPER 31: Statistics**

**Thursday 20 June 2024 – Afternoon**

**Question Booklet**

**DO NOT RETURN THIS BOOKLET WITH  
THE ANSWER BOOKLET.**

**YOU MUST HAVE**

**Mathematical Formulae and Statistical Tables (Green),  
calculator, writing and drawing equipment.**

**YOU WILL BE GIVEN**

**A separate Answer Booklet**

**A separate Diagram Booklet**

**Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for symbolic algebraic manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.**

**INSTRUCTIONS**

**Answer ALL questions and ensure that your answers to parts of questions are clearly labelled.**

**Answer the questions in the spaces provided in the Answer Booklet or in the separate Diagram Booklet – there may be more space than you need.**

**Do NOT write on this Question Booklet.**

**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 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.**

---

1. Xian rolls a fair die **10** times.

The random variable  **$X$**  represents the number of times the die lands on a six.

- (a) Using a suitable distribution for  **$X$** , find

(i)  **$P(X = 3)$**

(ii)  **$P(X < 3)$**

**(3 marks)**

Xian repeats this experiment each day for **60** days and records the number of days when  **$X = 3$**

- (b) Find the probability that there were at least **12** days when  **$X = 3$**

**(3 marks)**

- (c) Find an estimate for the total number of sixes that Xian will roll during these **60** days.

**(1 mark)**

1. continued.

- (d) Use a normal approximation to estimate the probability that Xian rolls a total of more than **95 sixes** during these **60** days.

(4 marks)

(Total for Question 1 is 11 marks)

---

2. Amar is studying the flight of a bird from its nest.

He measures the bird's height above the ground,  $h$  metres, at time  $t$  seconds for 10 values of  $t$

Amar finds the equation of the regression line for the data to be  $h = 38.6 - 1.28t$

- (a) Interpret the gradient of this line.  
(1 mark)
- (b) The product moment correlation coefficient between  $h$  and  $t$  is  $-0.510$

Test whether or not there is evidence of a negative correlation between the height above the ground and the time during the flight.

You should:

- state your hypotheses clearly
- use a 5% level of significance
- state the critical value used

(3 marks)

(continued on the next page)

2. continued.

- (c) Look at the diagram for Question 2 (c) in the separate Diagram Booklet. It shows the scatter diagram drawn by Jane for Amar's data.

With reference to the scatter diagram, state, giving a reason, whether or not the regression line  $h = 38.6 - 1.28t$  is an appropriate model for these data.  
(1 mark)

- (d) Jane suggests an improved model using the variable  $u = (t - k)^2$  where  $k$  is a constant.

She obtains the equation  $h = 38.1 - 0.78u$

Choose a suitable value for  $k$  to write Jane's improved model for  $h$  in terms of  $t$  only.  
(1 mark)

(Total for Question 2 is 6 marks)

---

**3. Ming is studying the large data set for Perth in 2015**

**He intended to use all the data available to find summary statistics for the Daily Mean Air Temperature,  $x$  °C.**

**Unfortunately, Ming selected an incorrect variable on the spreadsheet.**

**This incorrect variable gave a mean of 5.3 and a standard deviation of 12.4**

**(a) Using your knowledge of the large data set, suggest which variable Ming selected.**

**(1 mark)**

**(continued on the next page)**



3. continued.

(b) The correct values for the Daily Mean Air Temperature are summarised as

$$n = 184$$

$$\sum x = 2801.2$$

$$\sum x^2 = 44\,695.4$$

Calculate the mean and standard deviation for these data.

(3 marks)

(c) One of the months from the large data set for Perth in 2015 has

- mean  $\bar{x} = 19.4$
- standard deviation  $\sigma_x = 2.83$

for Daily Mean Air Temperature.

Suggest, giving a reason, a month these data may have come from.

(2 marks)

(Total for Question 3 is 6 marks)

---

4. The proportion of left-handed adults in a country is 10%

Freya believes that the proportion of left-handed adults under the age of 25 in this country is different from 10%

She takes a random sample of 40 adults under the age of 25 from this country to investigate her belief.

- (a) Find the critical region for a suitable test to assess Freya's belief.

You should

- state your hypotheses clearly
- use a 5% level of significance
- state the probability of rejection in each tail

(4 marks)

- (b) Write down the actual significance level of your test in part (a).

(1 mark)

(continued on the next page)

Turn over

**4. continued.**

**(c) In Freya's sample 7 adults were left-handed.**

**With reference to your answer in part (a) comment  
on Freya's belief.**

**(1 mark)**

**(Total for Question 4 is 6 marks)**

---

5. The records for a school athletics club show that the height,  $H$  metres, achieved by students in the high jump is normally distributed with mean  $1.4$  metres and standard deviation  $0.15$  metres.
- (a) Find the proportion of these students achieving a height of more than  $1.6$  metres.  
(1 mark)
- (b) The records also show that the time,  $T$  seconds, to run  $1500$  metres is normally distributed with mean  $330$  seconds and standard deviation  $26$  seconds.

The school's Head would like to use these distributions to estimate the proportion of students from the school athletics club who can jump higher than  $1.6$  metres AND can run  $1500$  metres in less than  $5$  minutes.

State a necessary assumption about  $H$  and  $T$  for the Head to calculate an estimate of this proportion.

(1 mark)

**5. continued.**

**(c) Find the Head's estimate of this proportion.**

**(3 marks)**

**(d) Students in the school athletics club also throw the discus.**

**The random variable  $D \sim N(\mu, \sigma^2)$  represents the distance, in metres, that a student can throw the discus.**

**Given that  $P(D < 16.3) = 0.30$  and**

**$P(D > 29.0) = 0.10$**

**calculate the value of  $\mu$  and the value of  $\sigma$**

**(5 marks)**

**(Total for Question 5 is 10 marks)**

---

6. Look at the diagram for Question 6 in the separate Diagram Booklet. It is a Venn diagram.

The Venn diagram, where  $p$ ,  $q$  and  $r$  are probabilities, shows the events  $A$ ,  $B$ ,  $C$  and  $D$  and associated probabilities.

- (a) State any pair of mutually exclusive events from  $A$ ,  $B$ ,  $C$  and  $D$ .  
(1 mark)

The events  $B$  and  $C$  are independent.

- (b) Find the value of  $p$   
(2 marks)

- (c) Find the greatest possible value of  $P(A | B')$   
(3 marks)

Given that  $P(B | A') = 0.5$

- (d) find the value of  $q$  and the value of  $r$   
(3 marks)

(continued on the next page)

Turn over

6. continued.

(e) Find  $P([A \cup B]' \cap C)$

(1 mark)

(f) Use set notation to write an expression for the event with probability  $p$

(1 mark)

(Total for Question 6 is 11 marks)

---

**TOTAL FOR STATISTICS IS 50 MARKS**

**END OF PAPER**

---