

Paper Reference 8MA0–21
Pearson Edexcel
Level 3 GCE

Mathematics
Advanced Subsidiary
Paper 21: Statistics

Wednesday 22 May 2019 – Morning

MATERIALS REQUIRED FOR EXAMINATION
Mathematical Formulae and Statistical Tables,
calculator

ITEMS INCLUDED WITH QUESTION PAPERS
Diagram Book
Answer Book

Q63360A

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.

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.

Answer ALL questions.

Write your answers in the Answer Book.

1. A sixth form college has **84** students in Year **12** and **56** students in Year **13**

The head teacher selects a stratified sample of **40** students, stratified by year group.

- (a) Describe how this sample could be taken.
(3 marks)

The head teacher is investigating the relationship between the amount of sleep, **S** hours, that each student had the night before they took an aptitude test and their performance in the test, **p** marks. For the sample of **40** students, he finds the equation of the regression line of **p** on **S** to be

$$p = 26 \cdot 1 + 5 \cdot 60s$$

- (b) With reference to this equation, describe the effect that an extra **0·5** hours of sleep may have, on average, on a student's performance in the aptitude test.
(1 mark)

(continued on the next page)

Turn over

1. continued.

(c) Describe one limitation of this regression model.

(1 mark)

(Total for Question 1 is 5 marks)

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

The Venn diagram shows three events, **A**, **B** and **C**, and their associated probabilities.

Events **B** and **C** are mutually exclusive.

Events **A** and **C** are independent.

Showing your working, find the value of **x**, the value of **y** and the value of **z**

(Total for Question 2 is 5 marks)

3. A fair **5**–sided spinner has sides numbered **1, 2, 3, 4** and **5**

The spinner is spun once and the score of the side it lands on is recorded.

- (a) Write down the name of the distribution that can be used to model the score of the side it lands on.

(1 mark)

The spinner is spun **28** times.

The random variable **X** represents the number of times the spinner lands on **2**

- (b) (i) Find the probability that the spinner lands on **2** at least **7** times.

(ii) Find **$P(4 \leq X < 8)$**

(5 marks)

(Total for Question 3 is 6 marks)

Turn over

- 4. Joshua is investigating the daily total rainfall in Hurn for May to October 2015**

Using the information from the large data set, Joshua wishes to calculate the mean of the daily total rainfall in Hurn for May to October 2015

- (a) Using your knowledge of the large data set, explain why Joshua needs to clean the data before calculating the mean.
(1 mark)**

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

Using the information from the large data set, he produces the grouped frequency table shown.

- (b) Use linear interpolation to calculate an estimate for the upper quartile of the daily total rainfall.
(2 marks)**

(continued on the next page)

Turn over

4. continued.

(c) Calculate an estimate for the standard deviation of the daily total rainfall in Hurn for May to October 2015

(2 marks)

(d) (i) State the assumption involved with using class midpoints to calculate an estimate of a mean from a grouped frequency table.

(ii) Using your knowledge of the large data set, explain why this assumption does not hold in this case.

(iii) State, giving a reason, whether you would expect the actual mean daily total rainfall in Hurn for May to October 2015 to be larger than, smaller than or the same as an estimate based on the grouped frequency table.

(3 marks)

(Total for Question 4 is 8 marks)

Turn over

- 5. Refer to the information for Question 5 in the Diagram Book.**

Past records show that 15% of customers at a shop buy chocolate.

The shopkeeper believes that moving the chocolate closer to the till will increase the proportion of customers buying chocolate.

After moving the chocolate closer to the till, a random sample of 30 customers is taken and 8 of them are found to have bought chocolate.

Julie carries out a hypothesis test, at the 5% level of significance, to test the shopkeeper's belief.

Julie's hypothesis test is shown in the Diagram Book.

- (a) Identify the first two errors that Julie has made in her hypothesis test.**

(2 marks)

(continued on the next page)

Turn over

5. continued.

(b) Explain whether or not these errors will affect the conclusion of her hypothesis test.

Give a reason for your answer.

(1 mark)

(c) Find, using a 5% level of significance, the critical region for a one-tailed test of the shopkeeper's belief.

The probability in the tail should be less than 0.05

(2 marks)

(d) Find the actual level of significance of this test.

(1 mark)

(Total for Question 5 is 6 marks)

TOTAL FOR STATISTICS IS 30 MARKS

END OF PAPER
