Rewarding Learning

**ADVANCED**
**General Certificate of Education**
**2010**

# Mathematics

## Assessment Unit S4

*assessing*

## Module S2: Statistics 2

**[AMS41]**

AMS41

**FRIDAY 18 JUNE, AFTERNOON**

**TIME**

1 hour 30 minutes.

**INSTRUCTIONS TO CANDIDATES**

Write your Centre Number and Candidate Number on the Answer Booklet provided.
Answer **all seven** questions.
Show clearly the full development of your answers.
Answers should be given to three significant figures unless otherwise stated.
You are permitted to use a graphic or scientific calculator in this paper.

**INFORMATION FOR CANDIDATES**

The total mark for this paper is 75
Figures in brackets printed down the right-hand side of pages indicate the marks awarded to each question or part question.
A copy of the **Mathematical Formulae and Tables booklet** is provided.
Throughout the paper the logarithmic notation used is $\ln z$ where it is noted that $\ln z \equiv \log_e z$

**BLANK PAGE**

**Answer all seven questions.**

**Show clearly the full development of your answers.**

**Answers should be given to three significant figures unless otherwise stated.**

**1** An experiment is carried out to explore the relationship between the heights of twenty-year-old men and the heights of their fathers. Eight pairs are measured. The results are given in **Table 1** below.

**Table 1**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Height of son, $x$ (cm) | 176 | 162 | 187 | 158 | 165 | 170 | 193 | 153 |
| Height of father, $y$ (cm) | 172 | 172 | 178 | 160 | 161 | 167 | 191 | 160 |

Summary values for these data are:

| $n$ | $\Sigma x$ | $\Sigma y$ | $\Sigma x^2$ | $\Sigma y^2$ | $\Sigma xy$ |
|---|---|---|---|---|---|
| 8 | 1364 | 1361 | 233 936 | 232 343 | 233 000 |

**(i)** Calculate the product–moment correlation coefficient for these data. [5]

**(ii)** Comment on the value obtained in part **(i)**. [1]

**2** In linear regression, one of the variables is called the *response* or *dependent* variable.

**(i)** Explain clearly what this means. Illustrate your answer with an example. [3]

Eamon carries out an experiment to explore the link between the percentage of carbon monoxide (CO) in the exhaust fumes of a car and the number of revolutions per minute (RPM) of the engine.

**(ii)** Give and briefly explain one modelling assumption for this experiment. [2]

Eamon's data is given in **Table 2** below.

**Table 2**

| RPM, $x$ | 1000 | 1500 | 2000 | 2500 | 3000 | 3500 |
|---|---|---|---|---|---|---|
| % CO, $y$ | 1.17 | 1.9 | 2.86 | 3.37 | 3.98 | 4.45 |

Summary values for these data are:

| $n$ | $\Sigma x$ | $\Sigma y$ | $\Sigma x^2$ | $\Sigma y^2$ | $\Sigma xy$ |
|---|---|---|---|---|---|
| 6 | 13 500 | 17.73 | 34 750 000 | 60.1583 | 45 680 |

**(iii)** Find the regression equation of percentage CO on RPM. [6]

**(iv)** If the car was running at 2200 RPM, estimate the percentage CO in the exhaust fumes. [2]

**3** Edith monitors the journey time, $x$ minutes, of trains. She suspects that the regular train from Belfast to Bangor does not take the thirty minutes that it should. Her sample data is summarised below.

$$n = 60 \qquad \Sigma x = 1721 \qquad \Sigma x^2 = 49\,441$$

**(i)** Calculate the mean and variance of the data. [3]

**(ii)** Test Edith's suspicion at 5% level. [10]

**4** Gareth is about to calculate a 95% confidence interval for the mean duration of telephone calls, duration $x$ mins, made from his office.

**(i)** Explain carefully what the expression "95% confidence interval" means. [2]

**(ii)** What modelling assumption is Gareth making in calculating the confidence interval? [1]

Gareth's data for his sample is summarised below.

$$n = 45 \qquad \Sigma x = 339 \qquad \Sigma x^2 = 2799$$

**(iii)** Find the mean and variance of the call durations. [3]

**(iv)** Find a 95% confidence interval for the mean call duration. [5]

**5** Ten Year 8 Mathematics students are selected to test the effectiveness of a computer program designed to improve performance in Mathematics. They begin by taking a Mathematics examination. Next they follow a three month course using the computer program. Finally they take a similar Mathematics examination at the end of the course. The program makers claim that the students' scores should improve by 10. The results of the examinations are given in **Table 3** below.

**Table 3**

| Student | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| Initial Score | 56 | 67 | 48 | 70 | 38 | 66 | 54 | 70 | 45 | 51 |
| Final Score | 63 | 75 | 61 | 74 | 51 | 75 | 67 | 73 | 53 | 62 |

Test the program makers' claim at 5% level. [15]

**6** A continuous random variable $X$ has mean 35 and variance 12
Sixty observations on $X$ were taken at random.
Find the probability that the sample mean is less than 34.7 [6]

**7** The independent random variables $X$ and $Y$ are such that $X \sim N(20, 6)$ and $Y \sim N(25, 4)$.

Find:

**(i)** $P(X + Y > 43)$; [4]

**(ii)** $P(3X < 2Y)$. [7]

**THIS IS THE END OF THE QUESTION PAPER**