

General Certificate of Education  
January 2004  
Advanced Level Examination



**MATHEMATICS (SPECIFICATION A)**  
**Unit Statistics 4**

**MAS4/W**

Thursday 29 January 2004 Morning Session

**In addition to this paper you will require:**

- an 8-page answer book;
- the AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed: 1 hour 20 minutes

**Instructions**

- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MAS4/W.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.
- Tie loosely any additional sheets you have used to the back of your answer book before handing it to the invigilator.

**Information**

- The maximum mark for this paper is 60.
- Mark allocations are shown in brackets.

**Advice**

- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

---

Answer **all** questions.

---

- 1 A market trader sells ball-point pens on his stall. He sells the pens for a different fixed price,  $x$  pence, in each of six weeks. He notes the number of pens,  $y$ , that he sells in each of these six weeks. The results are shown in the following table.

$x$	10	15	20	25	30	35
$y$	68	60	55	48	38	32

Calculate the equation of the least squares regression line of  $y$  on  $x$ . *(5 marks)*

- 2 Candidates sitting a multiple choice test are required to tick exactly one of five possible answers to each of the 20 questions. A particular candidate answers 7 questions correctly.

Stating null and alternative hypotheses, investigate, at the 5% level of significance, the claim that this candidate has randomly selected the answer to each question. *(5 marks)*

- 3 An experiment is performed and, for 6 pairs of data values  $(x, y)$ , the product moment correlation coefficient is calculated as  $-0.9824$ , correct to four significant figures.

(a) State what this value tells you about the 6 pairs of data values. *(1 mark)*

(b) You are given the following summarised data for 6 pairs of data values  $(w, y)$ .

$$\sum w = 91 \quad \sum w^2 = 2275 \quad \sum wy = 1812$$

$$\sum y = 190 \quad \sum y^2 = 7296$$

Calculate the value of the product moment correlation coefficient between  $w$  and  $y$ , giving your answer to four significant figures. *(5 marks)*

(c) Given that  $w = x^2$ , explain the implication of your answer in part (b). *(1 mark)*

- 4 Five years ago, a random sample of 500 adult members of the population of a certain town showed 160 in favour of a controlled parking zone (CPZ) in the town centre.

A recent survey of 500 adult members of the population showed 205 in favour of the CPZ.

- (a) Calculate an approximate 99% confidence interval for the difference between the proportions in favour of the CPZ now and five years ago. *(6 marks)*
- (b) The leader of the opposition on the local council claims that the support for the CPZ is no different now from what it was five years ago.

State, giving a reason, whether or not you agree with the leader of the opposition.

*(2 marks)*

- 5 A trainee clerk was asked to estimate the weight, in grams, of each of a random sample of seven packets. The table below shows the actual weight and the estimate given by the trainee.

Packet	A	B	C	D	E	F	G
Actual weight	140	210	630	320	160	700	450
Estimated weight	100	150	500	250	100	500	350

- (a) (i) Calculate the value of Spearman's rank correlation coefficient between the actual weight and the estimated weight. *(5 marks)*
- (ii) Comment on the claim that your value of Spearman's rank correlation coefficient suggests that the trainee is good at estimating the weights of packets. *(2 marks)*
- (b) Assuming that these data are a random sample from a distribution with Spearman's rank correlation coefficient  $\rho_s$ , investigate, at the 1% level of significance, the hypothesis that  $\rho_s > 0$ . *(4 marks)*

**TURN OVER FOR THE NEXT QUESTION**

6 A survey of a random sample of 200 final-year students at Camford University, who were seeking employment, found that 168 had a job one year after graduating.

- (a) (i) Find an approximate 95% confidence interval for the proportion,  $p$ , of final-year students who find a job within one year of graduating from Camford University. (5 marks)
- (ii) Similar surveys were carried out at 19 other universities. A 95% confidence interval for the proportion of final-year students who find a job within one year of graduating was calculated for each university.

Assuming that this proportion is the same at all the universities considered, state the expected number of the 20 confidence intervals that will enclose the proportion's true value. (1 mark)

- (b) The appointments officer at Camford University believes the value of  $p$  to be 0.9.

Investigate, at the 1% level of significance, whether the appointments officer is over-stating the value of  $p$ . (6 marks)

7 Two independent normal distributions have unknown means  $\mu_1$  and  $\mu_2$ , but known variances  $\sigma_1^2$  and  $\sigma_2^2$ , respectively.

- (a) Given that  $\bar{X}_1$  and  $\bar{X}_2$  denote the means of random samples of sizes  $n_1$  and  $n_2$  drawn from each distribution respectively, show that  $\bar{X}_1 - \bar{X}_2$  is an unbiased estimator of  $\mu_1 - \mu_2$ . (2 marks)
- (b) Find an expression for  $V$ , the variance of  $\bar{X}_1 - \bar{X}_2$ . (2 marks)
- (c) You are given that  $n_1 + n_2 = n$ , where  $n$  is fixed.

- (i) By finding  $\frac{dV}{dn_1}$ , show that a necessary condition for  $V$  to be a minimum is

$$n_1 : n_2 = \sigma_1 : \sigma_2. \quad (5 \text{ marks})$$

- (ii) Hence find values for  $n_1$  and  $n_2$  which minimise  $V$  when

$$n = 560 \quad \sigma_1^2 = 0.0025 \quad \sigma_2^2 = 0.0081. \quad (3 \text{ marks})$$

**END OF QUESTIONS**