

General Certificate of Education
June 2007
Advanced Level Examination



MATHEMATICS
Unit Statistics 4

MS04

Monday 18 June 2007 9.00 am to 10.30 am

For this paper you must have:

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

You may use a graphics calculator.

Time allowed: 1 hour 30 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 MS04.
- Answer **all** questions.
- Show all necessary working; 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.

Information

- The maximum mark for this paper is 75.
- The marks for questions are shown in brackets.

Advice

- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

- 1 The headteacher of a school believes that the standard deviation of the annual number of new pupils joining the school is 10.

A statistician on the staff collects the following data on the number of new pupils joining the school during each of a sample of ten years.

124 123 139 136 128 125 128 133 131 133

Investigate, at the 5% level of significance, the headteacher's belief. Assume that these data may be regarded as a random sample from a normal distribution. (8 marks)

- 2 The discrete random variable X has a geometric distribution with parameter p .

- (a) Given that the value of the mean is 4 times that of the variance, find the value of p . (3 marks)
- (b) Hence determine $P(X > 4 \mid X > 2)$. (4 marks)

- 3 The assessment of a physics course has two components: a written examination and a practical test. Each component has a maximum mark of 75. The marks achieved by 10 students in each component are shown in the table.

Student	A	B	C	D	E	F	G	H	I	J
Written Mark	35	47	54	55	43	48	41	59	47	31
Practical Mark	57	63	47	72	73	27	39	60	53	22

- (a) Investigate, using a paired t -test and the 5% level of significance, whether the mean mark in the written examination is less than that in the practical test. (10 marks)
- (b) State **two** assumptions that were necessary in order to carry out the test in part (a). (2 marks)

- 4 (a) A continuous random variable X has probability density function

$$f(x) = \lambda e^{-\lambda x} \text{ for } x \geq 0$$

Prove that the mean value of X is $\frac{1}{\lambda}$. (4 marks)

- (b) The lifetime of a component in a machine is T hours, where T has probability density function

$$f(t) = \frac{1}{a} e^{-\frac{t}{a}} \text{ for } t \geq 0$$

The mean lifetime of these components is known to be 62.5 hours.

- (i) Find the value of $\frac{1}{a}$. (2 marks)
- (ii) Calculate the probability that a component will last for at least 80 hours. (4 marks)
- (iii) Given that a component has lasted for 80 hours, find the probability that it will last for a further 20 hours. (3 marks)

- 5 One hundred 1-millilitre samples of water were taken at random and the number of bacteria in each sample was counted. The results are shown in the table.

Number of bacteria	0	1	2	3	4	5	6	7
Frequency	7	15	27	25	11	10	3	2

- (a) For these data, show that the mean number of bacteria per 1-millilitre of water is 2.7. (2 marks)
- (b) Hence, using a χ^2 goodness of fit test with the 10% level of significance, investigate whether the number of bacteria per 1-millilitre of water can be modelled by a Poisson distribution. (11 marks)

Turn over for the next question

- 6 A random variable X is distributed with mean μ and variance σ^2 . Three independent observations, X_1 , X_2 and X_3 , are taken on X .

The combined statistic

$$T = aX_1 + bX_2 + cX_3$$

where a , b and c are constants, is used as an estimator for μ .

- (a) Show that, if T is an unbiased estimator for μ , then $a + b + c = 1$. (3 marks)

- (b) Two unbiased estimators for μ are T_1 and T_2 , defined by

$$T_1 = \frac{1}{3}X_1 + \frac{1}{2}X_2 + \frac{1}{6}X_3$$

$$T_2 = \frac{2}{3}X_1 + \frac{3}{4}X_2 - \frac{5}{12}X_3$$

- (i) Calculate the relative efficiency of T_1 with respect to T_2 . (5 marks)
- (ii) With reference to your answer to part (b)(i), state, with a reason, which of T_1 and T_2 is the better unbiased estimator for μ . (2 marks)

- 7 A student at an agricultural college was asked to compare the variability of the weight, X grams, of eggs laid by free-range hens with the weight, Y grams, of eggs laid by battery hens.

The variables X and Y may be assumed to be normally distributed with variances σ_X^2 and σ_Y^2 respectively.

A random sample of 12 values of X resulted in $\sum(x - \bar{x})^2 = 761.2$, where \bar{x} denotes the sample mean.

A random sample of 10 values of Y resulted in $\sum(y - \bar{y})^2 = 386.1$, where \bar{y} denotes the sample mean.

- (a) Calculate unbiased estimates of σ_X^2 and σ_Y^2 . (2 marks)
- (b) (i) Hence determine a 90% confidence interval for the ratio $\frac{\sigma_X^2}{\sigma_Y^2}$. (8 marks)
- (ii) Comment on the suggestion that the weights of eggs laid by free-range hens are more variable than the weights of eggs laid by battery hens. (2 marks)

END OF QUESTIONS