General Certificate of Education
June 2006
Advanced Level Examination

MATHEMATICS **MS04**
Unit Statistics 4

Friday 23 June 2006 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 crossing times for a ferry travelling from Port P on the English coast to Port Q on the French coast were recorded on 8 occasions.

The times, correct to the nearest minute, were

95 92 102 90 81 84 109 108

The crossing times may be assumed to be a random sample from a normal distribution with standard deviation σ.

(a) Calculate a 95% confidence interval for σ^2. *(6 marks)*

(b) Comment on the suggestion that $\sigma = 6$. *(2 marks)*

2 The numbers of girls in 240 families, each having 3 children, are recorded below.

Number of girls	0	1	2	3
Number of families	36	96	83	25

(a) Test, at the 5% level of significance, the hypothesis that these data may be modelled by a binomial distribution with parameter $p = \frac{1}{2}$. *(8 marks)*

(b) (i) Explain how you would estimate p if it was not known to be $\frac{1}{2}$. *(2 marks)*

(ii) State the number of degrees of freedom that there would have been in your test in part (a) if p had needed to be estimated from the data. *(1 mark)*

3 (a) The time, in years, that a taxi driver keeps his taxi, before replacing it with a new one, can be modelled by an exponential distribution with parameter 0.2.

Find the probability that he keeps his taxi:

(i) for less than 2 years; *(2 marks)*

(ii) for more than 3 years. *(2 marks)*

(b) The continuous random variable X has an exponential distribution with probability density function $\mathrm{f}(x)$, where

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

Use integration to find:

(i) $\mathrm{E}(X)$; *(4 marks)*

(ii) the median value of X. *(3 marks)*

(c) Breakdowns occur on a stretch of road at a mean rate of 0.3 per day. The number of breakdowns follows a Poisson distribution.

Find, **in hours**:

(i) the mean time between breakdowns; *(1 mark)*

(ii) the median time between breakdowns. *(2 marks)*

4 The numbers of red blood cells, measured in millions per cubic millimetre of blood, for 10 women and 8 men were found to be as follows:

Women	5.05	3.98	4.73	5.36	4.92	5.44	4.04	4.40	4.15	5.33
Men	4.23	4.92	5.53	5.33	5.31	4.86	5.36	4.75		

Assume that these are independent random samples from normal populations.

(a) Show, at the 5% level of significance, that the hypothesis that the population variances are equal is accepted. *(8 marks)*

(b) Investigate, at the 5% level of significance, the hypothesis that the mean number of red blood cells is greater for men than for women. *(9 marks)*

Turn over for the next question

5 A random variable X has mean 2μ and variance 13, and an independent random variable Y has mean μ and variance 3. The random variable $aX + bY$ is an unbiased estimator of μ, where a and b are constants.

(a) Show that $2a + b = 1$. *(2 marks)*

(b) Show that $\text{Var}(aX + bY) = 3 - 12a + 25a^2$. *(3 marks)*

(c) Find values of a and b such that $aX + bY$ has minimum variance. *(3 marks)*

(d) A single observation is made on each of X and Y. The values observed are 15 and 10 respectively.

Obtain an unbiased estimate of μ which has minimum variance. *(2 marks)*

6 (a) Javinder is trying to start his old motorcycle which is known, on average, to start twice in every five attempts. It may be assumed that each attempt is independent of every other attempt, and that the probability of it starting on any attempt remains constant.

Calculate the probability that:

(i) his motorcycle will start on the third attempt; *(2 marks)*

(ii) it will take more than three attempts to start his motorcycle. *(2 marks)*

(b) The discrete random variable X has a geometric distribution with parameter p.

(i) Prove that $\text{E}(X) = \dfrac{1}{p}$. *(3 marks)*

(ii) Given that $\text{E}(X^2) = \dfrac{2-p}{p^2}$, show that $\text{Var}(X) = \dfrac{1-p}{p^2}$. *(2 marks)*

(c) Kylie's old car has a faulty starter motor. The number of attempts, Y, required to start her car may be assumed to follow a geometric distribution with parameter p, such that

$$\text{P}(Y = 1 \text{ or } 2) = 0.36$$

(i) Verify that the value of p is 0.2. *(1 mark)*

(ii) State the values of the mean, μ, and variance, σ^2, of Y. *(2 marks)*

(iii) Hence calculate $\text{P}\left(Y \leqslant \mu - \dfrac{\sigma}{\sqrt{5}}\right)$. *(3 marks)*

END OF QUESTIONS