Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
TOTAL	

AQA

General Certificate of Education
Advanced Level Examination
June 2015

Mathematics MS04

Unit Statistics 4

Wednesday 24 June 2015 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

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

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

Do not write outside the box

Answer **all** questions.

Answer each question in the space provided for that question.

1 The random variable X has a geometric distribution with parameter 0.25.

(a) Calculate $P(X = 4)$.

[1 mark]

(b) Calculate the smallest value of n such that $P(X = n) < 0.001$.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 1

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 1

2 Roberta, who works for her local council's trading-standards department, regularly buys, from her local market, bags of seeds and bags of nuts for wild birds.

She has concerns about the apparent large variability in the weights of the bags of seeds that she buys.

Roberta therefore records the weight, x grams, of each of 12 bags of seeds that she buys and finds that $\sum(x-\bar{x})^2 = 37\,100.25$.

The weights of bags of seeds may be modelled by a normal distribution with variance ${\sigma_S}^2$, and Roberta's 12 bags of seeds may be considered to be a random sample.

(a) Investigate, at the 5% level of significance, the hypothesis that $\sigma_S > 45$.

[6 marks]

(b) Roberta also records the weight, y grams, of each of 8 bags of nuts that she buys and finds that $\sum(y-\bar{y})^2 = 2033.50$.

The weights of bags of nuts may be modelled by a normal distribution with variance ${\sigma_N}^2$, and Roberta's 8 bags of nuts may be considered to be a random sample.

Test, at the 10% level of significance, the hypothesis that $\sigma_S = 2\sigma_N$.

[7 marks]

QUESTION PART REFERENCE	Answer space for question 2

QUESTION PART REFERENCE

Answer space for question 2

QUESTION PART REFERENCE

Answer space for question 2

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 2

Turn over ▶

3 The owner of a fish-and-chip shop buys washed potatoes from two suppliers, *Pots4U* and *WeRPots*. Each supplier delivers potatoes in 25 kg bags.

The owner weighs 15 bags of potatoes from *Pots4U* and 10 bags of potatoes from *WeRPots*. The results, in kilograms, are given in the table.

Pots4U	***WeRPots***
25.4	25.5
25.7	25.4
25.2	25.4
25.0	24.9
24.8	25.4
25.4	25.3
25.6	25.1
25.6	24.7
25.8	25.0
25.5	25.3
25.4	
25.2	
26.0	
24.9	
25.5	

The two samples of bags may each be assumed to be random, and the weights of potatoes in bags from the two suppliers may be modelled by normal distributions with the same variance.

(a) Construct a 95% confidence interval for the difference between the mean weight of bags of potatoes supplied by *Pots4U* and that of bags supplied by *WeRPots*.

[10 marks]

(b) What does your confidence interval reveal? Justify your answer.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

Do not write outside the box

QUESTION PART REFERENCE	**Answer space for question 3**

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 3

4 (a) The random variable X has a probability density function given by

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & x \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

(i) Use integration to find an expression for:

(A) $E(X)$;

(B) $P(X > x)$ for $x \geqslant 0$.

[5 marks]

(ii) Given that m denotes the median of X, evaluate $P(m < X \leqslant \mu)$.

[2 marks]

(b) During the manufacture of barbed wire, the length, X kilometres, between successive faults may be modelled by an exponential distribution with mean 2.

(i) Determine the probability that the length between successive faults is between 250 metres and 1250 metres.

[3 marks]

(ii) A farmer purchases 6 reels, each containing 250 metres of barbed wire.

Calculate the probability that at least 5 of the 6 reels contain wire with no faults.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

Turn over ▶

5 It is suggested that the difference, D minutes, between the time that a patient is actually seen by an osteopath and the patient's scheduled appointment time can be modelled by $D = 5 + X$, where X has the following probability density function.

$$f(x) = \begin{cases} \frac{1}{18}x^2 & 0 \leqslant x \leqslant 3 \\ \frac{1}{4}(5 - x) & 3 \leqslant x \leqslant 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Complete the table of **exact** probabilities, **Table 1**, shown on the opposite page.

[4 marks]

(b) The results of a random sample of 540 observations of D gave the frequencies shown in **Table 2**.

Table 2

D	0–5	5–6	6–7	7–8	8–9	9–10	>10
Number of appointments	3	11	63	182	222	53	6

Use a χ^2 goodness of fit test and the 5% level of significance to assess the suitability of the suggested model for D.

[8 marks]

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Table 1

D	0–5	5–6	6–7	7–8	8–9	9–10	>10
Probability	0			$\frac{19}{54}$		$\frac{1}{8}$	0

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Turn over ▶

Do not write outside the box

6 The independent random variables U and V have means μ and 2μ respectively, and the same variance σ^2.

The variable $\overline{U}$ denotes the mean of a random sample of $\boldsymbol{n}$ observations of U, and the variable $\overline{V}$ denotes the mean of a random sample of $\mathbf{2}\boldsymbol{n}$ observations of V.

(a) Two estimators suggested for μ are

$$X_1 = \frac{1}{3}(\overline{U} + \overline{V}) \qquad \text{and} \qquad X_2 = \frac{1}{2}\left(\overline{U} + \frac{\overline{V}}{2}\right)$$

(i) Show that X_1 and X_2 are both unbiased estimators for μ.

[3 marks]

(ii) Derive simplified expressions for each of $\mathrm{Var}(X_1)$ and $\mathrm{Var}(X_2)$.

[3 marks]

(iii) Calculate the efficiency of X_2 relative to X_1.

[2 marks]

(b) A third unbiased estimator suggested for μ is

$$Y = c\overline{U} + (1 - c)\frac{\overline{V}}{2}$$

where c is a constant chosen to minimise $\mathrm{Var}(Y)$.

(i) Show that the value of c is $\frac{1}{9}$, and hence find an expression for $\mathrm{Var}(Y)$.

[4 marks]

(ii) A random sample of 10 observations of U gave $\overline{u} = 4.8$, and an independent random sample of 20 observations of V gave $\overline{v} = 12.3$.

Given that U and V are each normally distributed with variance 16, use the distribution of Y to test, at the 1% level of significance, the hypothesis that $\mu = 5$.

[7 marks]

QUESTION PART REFERENCE	Answer space for question 6

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 6

Turn over ▶

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE	Answer space for question 6

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 6

END OF QUESTIONS