Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
TOTAL	

AQA

General Certificate of Education
Advanced Level Examination
June 2014

Mathematics MS03

Unit Statistics 3

Monday 23 June 2014 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.

Answer **all** questions.

Answer each question in the space provided for that question.

1 A hotel's management is concerned about the quality of the free pens that it provides in its meeting rooms.

The hotel's assistant manager tests a random sample of 200 such pens and finds that 23 of them fail to write immediately.

(a) Calculate an approximate 96% confidence interval for the proportion of pens that fail to write immediately.

[5 marks]

(b) The supplier of the pens to the hotel claims that at most 2 in 50 pens fail to write immediately.

Comment, with numerical justification, on the supplier's claim.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 1

QUESTION PART REFERENCE

Answer space for question 1

2 Each household within a district council's area has two types of wheelie-bin: a black one for general refuse and a green one for garden refuse. Each type of bin is emptied by the council fortnightly.

The weight, in kilograms, of refuse emptied from a black bin can be modelled by the random variable $B \sim \mathrm{N}(\mu_B,\ 0.5625)$.

The weight, in kilograms, of refuse emptied from a green bin can be modelled by the random variable $G \sim \mathrm{N}(\mu_G,\ 0.9025)$.

The mean weight of refuse emptied from a random sample of 20 black bins was $21.35\,\mathrm{kg}$. The mean weight of refuse emptied from an independent random sample of 15 green bins was $21.90\,\mathrm{kg}$.

Test, at the 5% level of significance, the hypothesis that $\mu_B = \mu_G$.

[6 marks]

QUESTION PART REFERENCE

Answer space for question 2

QUESTION PART REFERENCE

Answer space for question 2

3 An investigation was carried out into the type of vehicle being driven when its driver was caught speeding. The investigation was restricted to drivers who were caught speeding when driving vehicles with at least 4 wheels.

An analysis of the results showed that 65% were driving cars (C), 20% were driving vans (V) and 15% were driving lorries (L).

Of those driving cars, 30% were caught by fixed speed cameras (F), 55% were caught by mobile speed cameras (M) and 15% were caught by average speed cameras (A).

Of those driving vans, 35% were caught by fixed speed cameras (F), 45% were caught by mobile speed cameras (M) and 20% were caught by average speed cameras (A).

Of those driving lorries, 10% were caught by fixed speed cameras (F), 65% were caught by mobile speed cameras (M) and 25% were caught by average speed cameras (A).

(a) Represent this information by a tree diagram on which are shown labels and percentages or probabilities.

[3 marks]

(b) Hence, or otherwise, calculate the probability that a driver, selected at random from those caught speeding:

(i) was driving either a car or a lorry and was caught by a mobile speed camera;

(ii) was driving a lorry, given that the driver was caught by an average speed camera;

(iii) was **not** caught by a fixed speed camera, given that the driver was **not** driving a car.

[8 marks]

(c) Three drivers were selected at random from those caught speeding by **fixed speed cameras**.

Calculate the probability that they were driving three different types of vehicle.

[4 marks]

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 3

Turn over ▶

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

Turn over ▶

4 A sample of 50 male *Eastern Grey* kangaroos had a mean weight of $42.6\,\text{kg}$ and a standard deviation of $6.2\,\text{kg}$.

A sample of 50 male *Western Grey* kangaroos had a mean weight of $39.7\,\text{kg}$ and a standard deviation of $5.3\,\text{kg}$.

(a) Construct a 98% confidence interval for the difference between the mean weight of male *Eastern Grey* kangaroos and that of male *Western Grey* kangaroos.

[5 marks]

(b) (i) What assumption about the selection of each of the two samples was it necessary to make in order that the confidence interval constructed in part **(a)** was valid?

[1 mark]

(ii) Why was it **not** necessary to assume anything about the distributions of the weights of male kangaroos in order that the confidence interval constructed in part **(a)** was valid?

[2 marks]

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

5 The numbers of daily morning operations, X, and daily afternoon operations, Y, in an operating theatre of a small private hospital can be modelled by the following bivariate probability distribution.

		Number of morning operations (X)					
		2	**3**	**4**	**5**	**6**	**P($Y = y$)**
Number of afternoon operations (Y)	**3**	0.00	0.05	0.20	0.20	0.05	0.50
	4	0.00	0.15	0.10	0.05	0.00	0.30
	5	0.05	0.05	0.10	0.00	0.00	0.20
	P($X = x$)	0.05	0.25	0.40	0.25	0.05	1.00

(a) (i) State why $\mathrm{E}(X) = 4$ and show that $\mathrm{Var}(X) = 0.9$.

[4 marks]

(ii) Given that

$$\mathrm{E}(Y) = 3.7,\ \mathrm{Var}(Y) = 0.61 \text{ and } \mathrm{E}(XY) = 14.4$$

calculate values for $\mathrm{Cov}(X, Y)$ and ρ_{XY}.

[4 marks]

(b) Calculate values for the mean and the variance of:

(i) $T = X + Y$;

(ii) $D = X - Y$.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 5

QUESTION PART REFERENCE

Answer space for question 5

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Turn over ▶

6 *Population* A has a normal distribution with unknown mean μ_A and a variance of 18.8.

Population B has a normal distribution with unknown mean μ_B but with the same variance as *Population* A.

The random variables $\overline{X}_A$ and $\overline{X}_B$ denote the means of independent samples, each of size n, from populations A and B respectively.

(a) Find an expression, in terms of n, for $\mathrm{Var}(\overline{X}_A - \overline{X}_B)$.

[2 marks]

(b) Given that the width of a 99% confidence interval for $\mu_A - \mu_B$ is to be at most 5, calculate the minimum value for n.

[5 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 ▶

Do not write outside the box

7 (a) The random variable X has a Poisson distribution with parameter λ.

(i) Prove, from first principles, that $\mathrm{E}(X) = \lambda$.

[3 marks]

(ii) Given that $\mathrm{E}(X^2 - X) = \lambda^2$, deduce that $\mathrm{Var}(X) = \lambda$.

[1 mark]

(b) The number of faults in a 100-metre ball of nylon string may be modelled by a Poisson distribution with parameter λ.

(i) An analysis of one ball of string, selected at random, showed 15 faults.

Using an exact test, investigate the claim that $\lambda > 10$. Use the 5% level of significance.

[5 marks]

(ii) A subsequent analysis of a random sample of 20 balls of string showed a total of 241 faults.

(A) Using an approximate test, re-investigate the claim that $\lambda > 10$. Use the 5% level of significance.

[4 marks]

(B) Determine the critical value of the total number of faults for the test in part **(b)(ii)(A)**.

[3 marks]

(C) Given that, in fact, $\lambda = 12$, estimate the probability of a Type II error for a test of the claim that $\lambda > 10$ based upon a random sample of 20 balls of string and using the 5% level of significance.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 7

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 7

Turn over ▶

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

END OF QUESTIONS

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

There are no questions printed on this page

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED