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 2012

Mathematics   MS03

Unit Statistics 3

Friday 22 June 2012   1.30 pm to 3.00 pm

For this paper you must have:
• the blue AQA booklet of formulae and statistical tables.
You may use a graphics calculator.

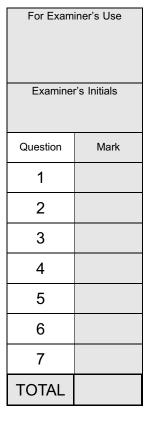## Time allowed
• 1 hour 30 minutes

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 wildlife expert measured the neck lengths, x metres, and the tail lengths, y metres, of a sample of 12 mature male giraffes as part of a study into their physical characteristics. The results are shown in the table.

x	1.62	1.81	1.75	1.59	1.66	1.61	1.73	1.81	1.88	1.72	1.62	1.60
y	2.33	2.48	2.40	2.31	2.37	2.29	2.47	2.46	2.51	2.34	2.44	2.46

(a) Calculate the value of the product moment correlation coefficient between x and y.

(2 marks)

(b) Investigate, at the 1% level of significance, the hypothesis that there is a positive correlation between the neck length and the tail length of mature male giraffes. The sample of measurements may be regarded as a random sample from a bivariate normal distribution.

(4 marks)

QUESTION PART REFERENCE

Answer space for question 1

QUESTION PART REFERENCE	Answer space for question 1

2 As part of a comparison of two varieties of cucumber, *Fanfare* and *Marketmore*, random samples of harvested cucumbers of each variety were selected and their lengths measured, in centimetres. The results are summarised in the table.

			Length (cm)	
		Sample size	Sample mean	Sample standard deviation
Cucumber variety	*Fanfare*	50	22.0	1.31
	Marketmore	75	21.6	0.702

(a) Test, at the 1% level of significance, the hypothesis that there is no difference between the mean length of harvested *Fanfare* cucumbers and that of harvested *Marketmore* cucumbers. *(6 marks)*

(b) In addition to length, name one other characteristic of cucumbers that could be used for comparative purposes. *(1 mark)*

QUESTION PART REFERENCE

Answer space for question 2

QUESTION PART REFERENCE	Answer space for question 2

3   A hotel has three types of room: double, twin and suite. The **percentage** of rooms in the hotel of each type is 40, 45 and 15 respectively.

Each room in the hotel may be occupied by 0, 1, 2, or 3 or more people.

The **proportional** occupancy of **each** type of room is shown in the table.

		Occupancy			
		0	**1**	**2**	**3 or more**
	Double	0.15	0.35	0.45	0.05
Room	**Twin**	0.05	0.55	0.30	0.10
	Suite	0.10	0.20	0.55	0.15

For example, the probability that, on a particular night, a double room has exactly 2 occupants is 0.45.

(a)   On a particular night, a room is selected at random. Find the probability that this room is:

(i)   an unoccupied suite; *(1 mark)*

(ii) occupied by 2 or more people; *(2 marks)*

(iii) unoccupied; *(2 marks)*

(iv) a double room, given that it is unoccupied; *(2 marks)*

(v)   a suite, given that it is occupied. *(3 marks)*

(b)   The hotel has a very large number of rooms from which, on a particular night, a random sample of 3 rooms is selected.

Given that all 3 of these rooms are occupied, calculate an estimate of the probability that each room is of a different type. *(4 marks)*

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

QUESTION PART REFERENCE	Answer space for question 3

QUESTION PART REFERENCE	Answer space for question 3

QUESTION PART REFERENCE	Answer space for question 3

4   The manager of a medical centre suspects that patients using repeat prescriptions were requesting, on average, more items during 2011 than during 2010.

The mean number of items on a repeat prescription during 2010 was 2.6 .

An analysis of a random sample of 250 repeat prescriptions during 2011 showed a total of 688 items requested.

The number of items requested on a repeat prescription may be modelled by a Poisson distribution.

Use a distributional approximation to investigate, at the 5% level of significance, the manager's suspicion.   *(6 marks)*

QUESTION PART REFERENCE	**Answer space for question 4**

QUESTION PART REFERENCE	Answer space for question 4

Turn over ▶

12

5 A random sample of 125 people was selected from a council's electoral roll. Of these, 68 were in favour of a proposed local building plan.

(a) Construct an approximate 98% confidence interval for the **percentage** of people on the council's electoral roll who were in favour of the proposal. *(6 marks)*

(b) Calculate, to the nearest 5, an estimate of the minimum sample size necessary in order that an approximate 98% confidence interval for the percentage of people on the council's electoral roll who were in favour of the proposal has a width of at most 10 per cent. *(4 marks)*

QUESTION PART REFERENCE	**Answer space for question 5**

QUESTION PART REFERENCE	Answer space for question 5

6   Alyssa lives in the country but works in a city centre.

Her journey to work each morning involves a car journey, a walk and wait, a train journey, and a walk.

Her car journey time, U minutes, from home to the village car park has a mean of 13 and a standard deviation of 3.

Her time, V minutes, to walk from the village car park to the village railway station and wait for a train to depart has a mean of 15 and a standard deviation of 6.

Her train journey time, W minutes, from the village railway station to the city centre railway station has a mean of 24 and a standard deviation of 4.

Her time, X minutes, to walk from the city centre railway station to her office has a mean of 9 and a standard deviation of 2.

The values of the product moment correlation coefficient for the above 4 variables are

$$\rho_{UV} = -0.6 \quad \text{and} \quad \rho_{UW} = \rho_{UX} = \rho_{VW} = \rho_{VX} = \rho_{WX} = 0$$

(a)   Determine values for the mean and the variance of:

    (i) $M = U + V$;   *(4 marks)*

    (ii) $D = W - 2U$;   *(3 marks)*

    (iii) $T = M + W + X$, given that $\rho_{MW} = \rho_{MX} = 0$.   *(2 marks)*

(b)   Assuming that the variables M, D and T are normally distributed, determine the probability that, on a particular morning:

    (i)   Alyssa's journey time from leaving home to leaving the village railway station is exactly 30 minutes;   *(1 mark)*

    (ii) Alyssa's train journey time is more than twice her car journey time;   *(3 marks)*

    (iii) Alyssa's total journey time is between 50 minutes and 70 minutes.   *(4 marks)*

QUESTION PART REFERENCE	**Answer space for question 6**

QUESTION PART REFERENCE	Answer space for question 6

P47644/Jun12/MS03

Turn over ▶

QUESTION PART REFERENCE	Answer space for question 6

QUESTION PART REFERENCE	Answer space for question 6

7 (a)   The random variable X has a binomial distribution with parameters n and p.

  (i) Prove, from first principles, that $E(X) = np$.   *(3 marks)*

  (ii) Hence, given that $E(X(X-1)) = n(n-1)p^2$, find, in terms of n and p, an expression for $\text{Var}(X)$.   *(2 marks)*

  (b)   The mode, m, of X is such that

$$P(X = m) \geqslant P(X = m - 1) \quad \text{and} \quad P(X = m) \geqslant P(X = m + 1)$$

  (i) Use the **first** inequality to show that

$$m \leqslant (n + 1)p$$ *(4 marks)*

  (ii) Given that the **second** inequality results in

$$m \geqslant (n + 1)p - 1$$

deduce that the distribution $B(10, 0.65)$ has one mode, and find the two values for the mode of the distribution $B(35, 0.5)$.   *(3 marks)*

  (c)   The random variable Y has a binomial distribution with parameters 4000 and 0.00095.

Use a distributional approximation to estimate $P(Y \leqslant k)$, where k denotes the mode of Y.   *(3 marks)*

QUESTION PART REFERENCE	**Answer space for question 7**

QUESTION PART REFERENCE	Answer space for question 7

P47644/Jun12/MS03

Turn over ▶

QUESTION PART REFERENCE	**Answer space for question 7**

P47644/Jun12/MS03

END OF QUESTIONS