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	

General Certificate of Education
Advanced Level Examination
June 2015

Mathematics MS03

Unit Statistics 3

Monday 22 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 A demographer measured the length of the right foot, x millimetres, and the length of the right hand, y millimetres, of each of a sample of 12 males aged between 19 years and 25 years. The results are given in the table.

x	215	221	228	235	241	248	258	261	272	278	283	296
y	166	174	171	179	187	179	183	180	192	202	187	208

$$S_{xx} = 7410 \qquad S_{yy} = 1642 \qquad S_{xy} = 3095$$

(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 length of the right foot and the length of the right hand of males aged between 19 years and 25 years. 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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 1

Do not write outside the box

2 Emilia runs an online perfume business from home. She believes that she receives more orders on Mondays than on Fridays.

She checked this during a period of 26 weeks and found that she received a total of 507 orders on the Mondays and a total of 416 orders on the Fridays.

The daily numbers of orders that Emilia receives may be modelled by independent Poisson distributions with means λ_{M} for Mondays and λ_{F} for Fridays.

(a) Construct an approximate 99% confidence interval for $\lambda_{\mathrm{M}} - \lambda_{\mathrm{F}}$.

[6 marks]

(b) Hence comment on Emilia's belief.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 2

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 2

Do not write outside the box

3 A particular brand of spread is produced in three varieties: standard, light and very light.

During a marketing campaign, the producer advertises that some cartons of spread contain coupons worth £1, £2 or £4.

For each variety of spread, the **proportion** of cartons containing coupons of each value is shown in the table.

	Variety		
	Standard	**Light**	**Very light**
No coupon	0.70	0.65	0.55
£1 coupon	0.20	0.25	0.30
£2 coupon	0.08	0.06	0.10
£4 coupon	0.02	0.04	0.05

For example, the probability that a carton of standard spread contains a coupon worth £2 is 0.08.

In a large batch of cartons, 55 per cent contain standard spread, 30 per cent contain light spread and 15 per cent contain very light spread.

(a) A carton of spread is selected at random from the batch. Find the probability that the carton:

(i) contains standard spread and a coupon worth £1;

(ii) does not contain a coupon;

(iii) contains light spread, given that it does not contain a coupon;

(iv) contains very light spread, given that it contains a coupon.

[8 marks]

(b) A random sample of 3 cartons is selected from the batch.

Given that all of these 3 cartons contain a coupon, find the probability that they each contain a different variety of spread.

[4 marks]

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 3

Do not write outside the box

4 (a) A large survey in the USA establishes that 60 per cent of its residents own a smartphone.

A survey of 250 UK residents reveals that 164 of them own a smartphone.

Assuming that these 250 UK residents may be regarded as a random sample, investigate the claim that the percentage of UK residents owning a smartphone is the same as that in the USA. Use the 5% level of significance.

[7 marks]

(b) A random sample of 40 residents in a market town reveals that 5 of them own a 4G mobile phone.

Use an exact test to investigate, at the 5% level of significance, the belief that fewer than 25 per cent of the town's residents own a 4G mobile phone.

[5 marks]

(c) A marketing company needs to estimate the proportion of residents in a large city who own a 4G mobile phone. It wishes to estimate this proportion to within 0.05 with a confidence of 98%.

Given that the proportion is known to be at most 30 per cent, estimate the sample size necessary in order to meet the company's need.

[5 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

Do not write outside the box

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

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

[3 marks]

(ii) Given that $\mathrm{E}(X(X-1)) = n(n-1)p^2$, find an expression for $\mathrm{Var}(X)$.

[2 marks]

(b) (i) The random variable Y has a binomial distribution with $\mathrm{E}(Y) = 3$ and $\mathrm{Var}(Y) = 2.985$.

Find values for n and p.

[3 marks]

(ii) The random variable U has $\mathrm{E}(U) = 5$ and $\mathrm{Var}(U) = 6.25$.

Show that U does not have a binomial distribution.

[2 marks]

(c) The random variable V has the distribution $\mathrm{Po}(5)$ and $W = 2V + 10$.

Show that $\mathrm{E}(W) = \mathrm{Var}(W)$ but that W does not have a Poisson distribution.

[3 marks]

(d) The probability that, in a particular country, a person has blood group AB negative is 0.2 per cent. A sample of 5000 people is selected.

Given that the sample may be assumed to be random, use a distributional approximation to estimate the probability that at least 6 people but at most 12 people have blood group AB negative.

[3 marks]

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

6 (a) The independent random variables S and L have means μ_S and μ_L respectively, and a common variance of σ^2.

The variable $\overline{S}$ denotes the mean of a random sample of n observations on S and the variable $\overline{L}$ denotes the mean of a random sample of n observations on L.

Find a simplified expression, in terms of σ^2, for the variance of $\overline{L} - 2\overline{S}$.

[3 marks]

(b) A machine fills both small bottles and large bottles with shower gel. It is known that the volume of shower gel delivered by the machine is normally distributed with a standard deviation of $8\,\text{ml}$.

(i) A random sample of 25 small bottles filled by the machine contained a mean volume of $\overline{s} = 258\,\text{ml}$ of shower gel.

An independent random sample of 25 large bottles filled by the machine contained a mean volume of $\overline{l} = 522\,\text{ml}$ of shower gel.

Investigate, at the 10% level of significance, the hypothesis that the mean volume of shower gel in a large bottle is more than **twice** that in a small bottle.

[7 marks]

(ii) Deduce that, for the test of the hypothesis in part **(b)(i)**, the critical value of $\overline{L} - 2\overline{S}$ is 4.585, correct to three decimal places.

[2 marks]

(iii) In fact, the mean volume of shower gel in a large bottle exceeds twice that in a small bottle by $10\,\text{ml}$.

Determine the probability of a Type II error for a test of the hypothesis in part **(b)(i)** at the 10% level of significance, based upon random samples of 25 small bottles and 25 large bottles.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 6

Do not write outside the box

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

Copyright © 2015 AQA and its licensors. All rights reserved.