AQA Maths Statistics 3

Past Paper Pack

2006-2016

General Certificate of Education
June 2006
Advanced Level Examination

ASSESSMENT and QUALIFICATIONS ALLIANCE

MATHEMATICS **MS03**
Unit Statistics 3

Wednesday 21 June 2006 1.30 pm to 3.00 pm

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 MS03.
- 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 A council claims that 80 per cent of households are generally satisfied with the services it provides.

A random sample of 250 households shows that 209 are generally satisfied with the council's provision of services.

(a) Construct an approximate 95% confidence interval for the proportion of households that are generally satisfied with the council's provision of services. *(6 marks)*

(b) Hence comment on the council's claim. *(2 marks)*

2 The table below shows the heart rates, x beats per minute, and the systolic blood pressures, y milligrams of mercury, of a random sample of 10 patients undergoing kidney dialysis.

Patient	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
x	83	86	88	92	94	98	101	111	115	121
y	157	172	161	154	171	169	179	180	192	182

(a) Calculate the value of the product moment correlation coefficient for these data. *(3 marks)*

(b) Assuming that these data come from a bivariate normal distribution, investigate, at the 1% level of significance, the claim that, for patients undergoing kidney dialysis, there is a positive correlation between heart rate and systolic blood pressure. *(4 marks)*

3 Each enquiry received by a business support unit is dealt with by Ewan, Fay or Gaby. The probabilities of them dealing with an enquiry are 0.2, 0.3 and 0.5 respectively.

Of enquiries dealt with by Ewan, 60% are answered immediately, 25% are answered later the same day and the remainder are answered at a later date.

Of enquiries dealt with by Fay, 75% are answered immediately, 15% are answered later the same day and the remainder are answered at a later date.

Of enquiries dealt with by Gaby, 90% are answered immediately and the remainder are answered at a later date.

(a) Determine the probability that an enquiry:

(i) is dealt with by Gaby and answered immediately; *(1 mark)*

(ii) is answered immediately; *(3 marks)*

(iii) is dealt with by Gaby, given that it is answered immediately. *(3 marks)*

(b) Determine the probability that an enquiry is dealt with by Ewan, given that it is answered later the same day. *(4 marks)*

4 The table below shows the probability distribution for the number of students, R, attending classes for a particular mathematics module.

r	6	7	8
$\mathbf{P}(R = r)$	0.1	0.6	0.3

(a) Find values for $\mathrm{E}(R)$ and $\mathrm{Var}(R)$. *(4 marks)*

(b) The number of students, S, attending classes for a different mathematics module is such that

$$\mathrm{E}(S) = 10.9, \quad \mathrm{Var}(S) = 1.69 \quad \text{and} \quad \rho_{RS} = \frac{2}{3}$$

Find values for the mean and variance of:

(i) $T = R + S$; *(4 marks)*

(ii) $D = S - R$. *(2 marks)*

Turn over for the next question

5 The number of letters per week received at home by Rosa may be modelled by a Poisson distribution with parameter 12.25.

(a) Using a normal approximation, estimate the probability that, during a 4-week period, Rosa receives at home at least 42 letters but at most 54 letters. *(5 marks)*

(b) Rosa also receives letters at work. During a 16-week period, she receives at work a total of 248 letters.

(i) Assuming that the number of letters received at work by Rosa may also be modelled by a Poisson distribution, calculate a 98% confidence interval for the average number of letters per week received at work by Rosa. *(5 marks)*

(ii) Hence comment on Rosa's belief that she receives, on average, fewer letters at home than at work. *(2 marks)*

6 The random variable X has a Poisson distribution with parameter λ.

(a) Prove that $\mathrm{E}(X) = \lambda$. *(3 marks)*

(b) By first proving that $\mathrm{E}(X(X-1)) = \lambda^2$, or otherwise, prove that $\mathrm{Var}(X) = \lambda$. *(5 marks)*

7 A shop sells cooked chickens in two sizes: medium and large.

The weights, X grams, of medium chickens may be assumed to be normally distributed with mean μ_X and standard deviation 45.

The weights, Y grams, of large chickens may be assumed to be normally distributed with mean μ_Y and standard deviation 65.

A random sample of 20 medium chickens had a mean weight, $\bar{x}$ grams, of 936.

A random sample of 10 large chickens had the following weights in grams:

1165 1202 1077 1144 1195 1275 1136 1215 1233 1288

(a) Calculate the mean weight, $\bar{y}$ grams, of this sample of large chickens. *(1 mark)*

(b) Hence investigate, at the 1% level of significance, the claim that the mean weight of large chickens exceeds that of medium chickens by more than 200 grams. *(8 marks)*

(c) (i) Deduce that, for your test in part (b), the critical value of $(\bar{y} - \bar{x})$ is 253.24, correct to two decimal places. *(2 marks)*

(ii) Hence determine the power of your test in part (b), given that $\mu_Y - \mu_X = 275$. *(5 marks)*

(iii) Interpret, in the context of this question, the value that you obtained in part (c)(ii). *(3 marks)*

END OF QUESTIONS

General Certificate of Education
June 2007
Advanced Level Examination

MATHEMATICS
Unit Statistics 3

MS03

ASSESSMENT and QUALIFICATIONS ALLIANCE

Monday 11 June 2007 1.30 pm to 3.00 pm

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 MS03.
- 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 As part of an investigation into the starting salaries of graduates in a European country, the following information was collected.

		Starting salary (€)	
	Sample size	**Sample mean**	**Sample standard deviation**
Science graduates	175	19 268	7321
Arts graduates	225	17 896	8205

(a) Stating a necessary assumption about the samples, construct a 98% confidence interval for the difference between the mean starting salary of science graduates and that of arts graduates. *(6 marks)*

(b) What can be concluded from your confidence interval? *(2 marks)*

2 A hill-top monument can be visited by one of three routes: road, funicular railway or cable car. The percentages of visitors using these routes are 25, 35 and 40 respectively.

The age distribution, in percentages, of visitors using **each** route is shown in the table. For example, 15 per cent of visitors using the road were under 18.

		Percentage of visitors using		
		Road	**Funicular railway**	**Cable car**
Age (years)	**Under 18**	15	25	10
	18 to 64	80	60	55
	Over 64	5	15	35

Calculate the probability that a randomly selected visitor:

(a) who used the road is aged 18 or over; *(1 mark)*

(b) is aged between 18 and 64; *(3 marks)*

(c) used the funicular railway and is aged over 64; *(2 marks)*

(d) used the funicular railway, given that the visitor is aged over 64. *(5 marks)*

3 Kutz and Styler are two unisex hair salons. An analysis of a random sample of 150 customers at Kutz shows that 28 per cent are male. An analysis of an independent random sample of 250 customers at Styler shows that 34 per cent are male.

(a) Test, at the 5% level of significance, the hypothesis that there is no difference between the proportion of male customers at Kutz and that at Styler. *(9 marks)*

(b) State, with a reason, the probability of making a Type I error in the test in part (a) if, in fact, the actual difference between the two proportions is 0.05. *(2 marks)*

4 A machine is used to fill 5-litre plastic containers with vinegar. The volume, in litres, of vinegar in a container filled by the machine may be assumed to be normally distributed with mean μ and standard deviation 0.08.

A quality control inspector requires a 99% confidence interval for μ to be constructed such that it has a width of at most 0.05 litres.

Calculate, to the nearest 5, the sample size necessary in order to achieve the inspector's requirement. *(6 marks)*

5 The duration, X minutes, of a timetabled 1-hour lesson may be assumed to be normally distributed with mean 54 and standard deviation 2.

The duration, Y minutes, of a timetabled $1\frac{1}{2}$-hour lesson may be assumed to be normally distributed with mean 83 and standard deviation 3.

Assuming the durations of lessons to be independent, determine the probability that the total duration of a random sample of three 1-hour lessons is less than the total duration of a random sample of two $1\frac{1}{2}$-hour lessons. *(7 marks)*

Turn over for the next question

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

(i) Prove that $\mathrm{E}(X) = np$. *(4 marks)*

(ii) Given that $\mathrm{E}(X^2) - \mathrm{E}(X) = n(n-1)p^2$, show that $\mathrm{Var}(X) = np(1-p)$. *(3 marks)*

(iii) Given that X is found to have a mean of 3 and a variance of 2.97, find values for n and p. *(3 marks)*

(iv) Hence use a distributional approximation to estimate $\mathrm{P}(X > 2)$. *(3 marks)*

(b) Dressher is a nationwide chain of stores selling women's clothes. It claims that the probability that a customer who buys clothes from its stores uses a Dressher store card is 0.45.

Assuming this claim to be correct, use a distributional approximation to estimate the probability that, in a random sample of 500 customers who buy clothes from Dressher stores, at least half of them use a Dressher store card. *(7 marks)*

7 In a town, the total number, R, of houses sold during a week by estate agents may be modelled by a Poisson distribution with a mean of 13.

A new housing development is completed in the town. During the first week in which houses on this development are offered for sale by the developer, the estate agents sell a total of 10 houses.

(a) Using the 10% level of significance, investigate whether the offer for sale of houses by the developer has resulted in a reduction in the mean value of R. *(6 marks)*

(b) Determine, for your test in part (a), the critical region for R. *(2 marks)*

(c) Assuming that the offer for sale of houses on the new housing development has reduced the mean value of R to 6.5, determine, for a test at the 10% level of significance, the probability of a Type II error. *(4 marks)*

END OF QUESTIONS

General Certificate of Education
June 2008
Advanced Level Examination

ASSESSMENT and QUALIFICATIONS ALLIANCE

MATHEMATICS **MS03**
Unit Statistics 3

Friday 23 May 2008 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 black ink or black 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 MS03.
- 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 best performances of a random sample of 20 junior athletes in the long jump, x metres, and in the high jump, y metres, were recorded. The following statistics were calculated from the results.

$$S_{xx} = 7.0036 \quad S_{yy} = 0.8464 \quad S_{xy} = 1.3781$$

(a) Calculate the value of the product moment correlation coefficient between x and y. *(2 marks)*

(b) Assuming that these data come from a bivariate normal distribution, investigate, at the 1% level of significance, the claim that for junior athletes there is a positive correlation between x and y. *(4 marks)*

(c) Interpret your conclusion in the context of this question. *(1 mark)*

2 A survey of a random sample of 200 passengers on UK internal flights revealed that 132 of them were on business trips.

(a) Construct an approximate 98% confidence interval for the proportion of passengers on UK internal flights that are on business trips. *(6 marks)*

(b) Hence comment on the claim that more than 60 per cent of passengers on UK internal flights are on business trips. *(2 marks)*

3 Pitted black olives in brine are sold in jars labelled "340 grams net weight". Two machines, A and B, independently fill these jars with olives before the brine is added.

The weight, X grams, of olives delivered by machine A may be modelled by a normal distribution with mean μ_X and standard deviation 4.5.

The weight, Y grams, of olives delivered by machine B may be modelled by a normal distribution with mean μ_Y and standard deviation 5.7.

The mean weight of olives from a random sample of 10 jars filled by machine A is found to be 157 grams, whereas that from a random sample of 15 jars filled by machine B is found to be 162 grams.

Test, at the 1% level of significance, the hypothesis that $\mu_X = \mu_Y$. *(6 marks)*

4 A manufacturer produces three models of washing machine: basic, standard and deluxe. An analysis of warranty records shows that 25% of faults are on basic machines, 60% are on standard machines and 15% are on deluxe machines.

For basic machines, 30% of faults reported during the warranty period are electrical, 50% are mechanical and 20% are water-related.

For standard machines, 40% of faults reported during the warranty period are electrical, 45% are mechanical and 15% are water-related.

For deluxe machines, 55% of faults reported during the warranty period are electrical, 35% are mechanical and 10% are water-related.

(a) Draw a tree diagram to represent the above information. *(3 marks)*

(b) Hence, or otherwise, determine the probability that a fault reported during the warranty period:

(i) is electrical; *(2 marks)*

(ii) is on a deluxe machine, given that it is electrical. *(2 marks)*

(c) A random sample of 10 electrical faults reported during the warranty period is selected. Calculate the probability that exactly 4 of them are on deluxe machines. *(3 marks)*

5 The daily number of emergency calls received from district A may be modelled by a Poisson distribution with a mean of λ_A.

The daily number of emergency calls received from district B may be modelled by a Poisson distribution with a mean of λ_B.

During a period of 184 days, the number of emergency calls received from district A was 3312, whilst the number received from district B was 2760.

(a) Construct an approximate 95% confidence interval for $\lambda_A - \lambda_B$. *(6 marks)*

(b) State one assumption that is necessary in order to construct the confidence interval in part (a). *(1 mark)*

6 An aircraft, based at airport A, flies regularly to and from airport B.

The aircraft's flying time, X minutes, from A to B has a mean of 128 and a variance of 50.

The aircraft's flying time, Y minutes, on the return flight from B to A is such that

$$\mathrm{E}(Y) = 112, \quad \mathrm{Var}(Y) = 50 \quad \text{and} \quad \rho_{XY} = -0.4$$

(a) Given that $F = X + Y$:

(i) find the mean of F;

(ii) show that the variance of F is 60. *(4 marks)*

(b) At airport B, the stopover time, S minutes, is independent of F and has a mean of 75 and a variance of 36.

Find values for the mean and the variance of:

(i) $T = F + S$; *(2 marks)*

(ii) $M = F - 3S$. *(3 marks)*

(c) Hence, assuming that T and M are normally distributed, determine the probability that, on a particular round trip of the aircraft from A to B and back to A:

(i) the time from leaving A to returning to A exceeds 300 minutes; *(3 marks)*

(ii) the stopover time is greater than one third of the total flying time. *(6 marks)*

7 (a) The random variable X has a Poisson distribution with $\mathrm{E}(X) = \lambda$.

(i) Prove, from first principles, that $\mathrm{E}(X(X-1)) = \lambda^2$. *(4 marks)*

(ii) Hence deduce that $\mathrm{Var}(X) = \lambda$. *(2 marks)*

(b) The independent Poisson random variables X_1 and X_2 are such that $\mathrm{E}(X_1) = 5$ and $\mathrm{E}(X_2) = 2$.

The random variables D and F are defined by

$$D = X_1 - X_2 \quad \text{and} \quad F = 2X_1 + 10$$

(i) Determine the mean and the variance of D. *(2 marks)*

(ii) Determine the mean and the variance of F. *(3 marks)*

(iii) For **each** of the variables D and F, give a reason why the distribution is **not** Poisson. *(2 marks)*

(c) The daily number of black printer cartridges sold by a shop may be modelled by a Poisson distribution with a mean of 5.

Independently, the daily number of colour printer cartridges sold by the same shop may be modelled by a Poisson distribution with a mean of 2.

Use a distributional approximation to estimate the probability that the total number of black and colour printer cartridges sold by the shop during a 4-week period (24 days) exceeds 175. *(6 marks)*

END OF QUESTIONS

There are no questions printed on this page

There are no questions printed on this page

There are no questions printed on this page

General Certificate of Education
June 2009
Advanced Level Examination

MATHEMATICS MS03
Unit Statistics 3

Friday 19 June 2009 1.30 pm to 3.00 pm

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 black ink or black 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 MS03.
- 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 An analysis of a random sample of 150 urban dwellings for sale showed that 102 are semi-detached.

An analysis of an independent random sample of 80 rural dwellings for sale showed that 36 are semi-detached.

(a) Construct an approximate 99% confidence interval for the difference between the proportion of urban dwellings for sale that are semi-detached and the proportion of rural dwellings for sale that are semi-detached. *(6 marks)*

(b) Hence comment on the claim that there is no difference between these two proportions. *(2 marks)*

2 A hotel chain has hotels in three types of location: city, coastal and country. The percentages of the chain's reservations for each of these locations are 30, 55 and 15 respectively.

Each of the chain's hotels offers three types of reservation: Bed & Breakfast, Half Board and Full Board.

The percentages of these types of reservation for **each** of the three types of location are shown in the table.

		Type of location		
		City	**Coastal**	**Country**
Type of reservation	**Bed & Breakfast**	80	10	30
	Half Board	15	65	50
	Full Board	5	25	20

For example, 80 per cent of reservations for hotels in city locations are for Bed & Breakfast.

(a) For a reservation selected at random:

(i) show that the probability that it is for Bed & Breakfast is 0.34; *(2 marks)*

(ii) calculate the probability that it is for Half Board in a hotel in a coastal location; *(2 marks)*

(iii) calculate the probability that it is for a hotel in a coastal location, given that it is for Half Board. *(4 marks)*

(b) A random sample of 3 reservations for Half Board is selected.

Calculate the probability that these 3 reservations are for hotels in different types of location. *(5 marks)*

3 The proportion, p, of an island's population with blood type A Rh^+ is believed to be approximately 0.35.

A medical organisation, requiring a more accurate estimate, specifies that a 98% confidence interval for p should have a width of at most 0.1.

Calculate, to the nearest 10, an estimate of the minimum sample size necessary in order to achieve the organisation's requirement. *(6 marks)*

4 Holly, a horticultural researcher, believes that the mean height of stems on Tahiti daffodils exceeds that on Jetfire daffodils by more than 15 cm.

She measures the heights, x centimetres, of stems on a random sample of 65 Tahiti daffodils and finds that their mean, $\bar{x}$, is 40.7 and that their standard deviation, s_x, is 3.4.

She also measures the heights, y centimetres, of stems on a random sample of 75 Jetfire daffodils and finds that their mean, $\bar{y}$, is 24.4 and that their standard deviation, s_y, is 2.8.

Investigate, at the 1% level of significance, Holly's belief. *(8 marks)*

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

(a) Given that

$$\mathrm{E}(X) = np \quad \text{and} \quad \mathrm{E}(X(X-1)) = n(n-1)p^2$$

find an expression for $\mathrm{Var}(X)$. *(3 marks)*

(b) Given that X has a mean of 36 and a standard deviation of 4.8:

(i) find values for n and p; *(3 marks)*

(ii) use a distributional approximation to estimate $\mathrm{P}(30 < X < 40)$. *(4 marks)*

6 The table shows the probability distribution for the number of weekday (Monday to Friday) morning newspapers, X, purchased by the Reed household per week.

x	0	1	2	3	4	5
$\mathbf{P}(X=x)$	0.16	0.15	0.25	0.25	0.15	0.04

(a) Find values for $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. *(3 marks)*

(b) The number of weekday (Monday to Friday) evening newspapers, Y, purchased by the same household per week is such that

$$\mathrm{E}(Y) = 2.0, \quad \mathrm{Var}(Y) = 1.5 \quad \text{and} \quad \mathrm{Cov}(X, Y) = -0.43$$

Find values for the mean and variance of:

(i) $S = X + Y$;

(ii) $D = X - Y$. *(5 marks)*

(c) The total cost per week, L, of the Reed household's weekday morning and evening newspapers may be assumed to be normally distributed with a mean of £2.31 and a standard deviation of £0.89.

The total cost per week, M, of the household's weekend (Saturday and Sunday) newspapers may be assumed to be independent of L and normally distributed with a mean of £2.04 and a standard deviation of £0.43.

Determine the probability that the total cost per week of the Reed household's newspapers is more than £5. *(5 marks)*

7 The daily number of customers visiting a small arts and crafts shop may be modelled by a Poisson distribution with a mean of 24.

(a) Using a distributional approximation, estimate the probability that there was a total of at most 150 customers visiting the shop during a given 6-day period. *(5 marks)*

(b) The shop offers a picture framing service. The daily number of requests, Y, for this service may be assumed to have a Poisson distribution.

Prior to the shop advertising this service in the local free newspaper, the mean value of Y was 2. Following the advertisement, the shop received a total of 17 requests for the service during a period of 5 days.

(i) Using a Poisson distribution, carry out a test, at the 10% level of significance, to investigate the claim that the advertisement increased the mean daily number of requests for the shop's picture framing service. *(5 marks)*

(ii) Determine the critical value of Y for your test in part (b)(i). *(3 marks)*

(iii) Hence, assuming that the advertisement increased the mean value of Y to 3, determine the power of your test in part (b)(i). *(4 marks)*

END OF QUESTIONS

There are no questions printed on this page

There are no questions printed on this page

There are no questions printed on this page

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	

General Certificate of Education
Advanced Level Examination
June 2010

Mathematics MS03

Unit Statistics 3

Tuesday 22 June 2010 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 the questions in the spaces provided. 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.

Answer **all** questions in the spaces provided.

1 Ffion, as part of her research project, measured the stem length and the cap diameter of each of a random sample of 24 matsutake mushrooms. Using these measurements, she calculated the value of the product moment correlation coefficient to be 0.336, correct to three significant figures.

Assuming that her measurements came from a bivariate normal distribution, test, at the 5% level of significance, the hypothesis that there is no correlation between the stem length and the cap diameter of matsutake mushrooms. *(4 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

2 Rodney and Derrick, two independent fruit and vegetable market stallholders, sell punnets of locally-grown raspberries from their stalls during June and July.

The following information, based on independent random samples, was collected as part of an investigation by Trading Standards Officers.

			Weight of raspberries in a punnet (grams)	
		Sample size	**Sample mean**	**Sample standard deviation, *s***
Stallholder	**Rodney**	50	225	5
	Derrick	75	219	8

(a) Construct a 99% confidence interval for the difference between the mean weight of raspberries in a punnet sold by Rodney and the mean weight of raspberries in a punnet sold by Derrick. *(5 marks)*

(b) What can be concluded from your confidence interval? *(2 marks)*

(c) In addition to weight, state one other factor that may influence whether customers buy raspberries from Rodney or from Derrick. *(1 mark)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

3 The weekly number of hits, S, on Sam's website may be modelled by a Poisson distribution with parameter λ_S. The weekly number of hits, T, on Tina's website may be modelled by a Poisson distribution with parameter λ_T.

During a period of 40 weeks, the number of hits on Sam's website was 940.

During a period of 60 weeks, the number of hits on Tina's website was 1560.

Assuming that S and T are independent random variables, investigate, at the 2% level of significance, Tina's claim that the mean weekly number of hits on her website is greater than that on Sam's website. *(7 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

4 It is proposed to introduce, for all males at age 60, screening tests, A and B, for a certain disease.

Test B is administered only when the result of Test A is inconclusive.

It is known that 10% of 60-year-old men suffer from the disease.

For those 60-year-old men suffering from the disease:

- Test A is known to give a positive result, indicating a presence of the disease, in 90% of cases, a negative result in 2% of cases and a requirement for the administration of Test B in 8% of cases;
- Test B is known to give a positive result in 98% of cases and a negative result in 2% of cases.

For those 60-year-old men not suffering from the disease:

- Test A is known to give a positive result in 1% of cases, a negative result in 80% of cases and a requirement for the administration of Test B in 19% of cases;
- Test B is known to give a positive result in 1% of cases and a negative result in 99% of cases.

(a) Draw a tree diagram to represent the above information. *(4 marks)*

(b) (i) Hence, or otherwise, determine the probability that:

(A) a 60-year-old man, suffering from the disease, tests negative;

(B) a 60-year-old man, not suffering from the disease, tests positive. *(2 marks)*

(ii) A random sample of ten thousand 60-year-old men is given the screening tests. Calculate, to the nearest 10, the number who you would expect to be given an **incorrect** diagnosis. *(2 marks)*

(c) Determine the probability that:

(i) a 60-year-old man suffers from the disease given that the tests provide a positive result;

(ii) a 60-year-old man does not suffer from the disease given that the tests provide a negative result. *(5 marks)*

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

5 In the manufacture of desk drawer fronts, a machine cuts sheets of veneered chipboard into rectangular pieces of width W millimetres and height H millimetres. The 4 edges of each of these pieces are then covered with matching veneered tape.

The distributions of W and H are such that

$$\mathrm{E}(W) = 350 \qquad \mathrm{Var}(W) = 5 \qquad \mathrm{E}(H) = 210 \qquad \mathrm{Var}(H) = 4 \qquad \rho_{WH} = 0.75$$

(a) Calculate the mean and the variance of the length of tape, $T = 2W + 2H$, needed for the edges of a drawer front. *(5 marks)*

(b) A desk has 4 such drawers whose sizes may be assumed to be independent.

Given that T may be assumed to be normally distributed, determine the probability that the total length of tape needed for the edges of the desk's 4 drawer fronts does not exceed 4.5 metres. *(5 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

6 (a) A district council claimed that more than 80 per cent of the complaints that it received about the delivery of its services were answered to the satisfaction of complainants before reaching formal status.

An analysis of a random sample of 175 complaints revealed that 28 reached formal status.

(i) Construct an approximate 95% confidence interval for the proportion of complaints that reach formal status. *(5 marks)*

(ii) Hence comment on the council's claim. *(2 marks)*

(b) The district council also claimed that less than 40 per cent of all formal complaints were due to a failing in the delivery of its services.

An analysis of the 50 formal complaints received during 2007/08 showed that 16 were due to a failing in the delivery of its services.

(i) Using an exact test, investigate the council's claim at the 10% level of significance. The 50 formal complaints received during 2007/08 may be assumed to be a random sample. *(5 marks)*

(ii) Determine the critical value for your test in part **(b)(i)**. *(2 marks)*

(iii) In fact, only 25 per cent of all formal complaints were due to a failing in the delivery of the council's services.

Determine the probability of a Type II error for a test of the council's claim at the 10% level of significance and based on the analysis of a random sample of 50 formal complaints. *(4 marks)*

QUESTION PART REFERENCE	
	.....
	.....
	.....
	.....
	.....
	.....
	.....
	.....
	.....

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

QUESTION PART REFERENCE

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

(a) (i) Prove, from first principles, that $\mathrm{E}(X) = \lambda$. *(3 marks)*

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

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

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

(i) Show that $\lambda - 1 \leqslant m \leqslant \lambda$. *(3 marks)*

(ii) Given that $\lambda = 4.9$, determine $\mathrm{P}(X = m)$. *(2 marks)*

(c) The random variable Y has a Poisson distribution with mode d and standard deviation 15.5.

Use a distributional approximation to estimate $\mathrm{P}(Y \geqslant d)$. *(5 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

END OF QUESTIONS

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	
8	
TOTAL	

General Certificate of Education
Advanced Level Examination
June 2011

Mathematics MS03

Unit Statistics 3

Wednesday 22 June 2011 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 the questions in the spaces provided. 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.

Answer **all** questions in the spaces provided.

1 A consumer report claimed that more than 25 per cent of visitors to a theme park were dissatisfied with the catering facilities provided.

In a survey, 375 visitors who had used the catering facilities were interviewed independently, and 108 of them stated that they were dissatisfied with the catering facilities provided.

(a) Test, at the 2% level of significance, the consumer report's claim. *(6 marks)*

(b) State an assumption about the 375 visitors that was necessary in order for the hypothesis test in part **(a)** to be valid. *(1 mark)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

2 The number of emergency calls received by a fire station may be modelled by a Poisson distribution.

During a given period of 13 weeks, the station received a total of 108 emergency calls.

(a) Construct an approximate 98% confidence interval for the average **weekly** number of emergency calls received by the station. *(5 marks)*

(b) Hence comment on the station officer's claim that the station receives an average of one emergency call **per day**. *(2 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

3 An IT help desk has three telephone stations: Alpha, Beta and Gamma. Each of these stations deals only with telephone enquiries.

The probability that an enquiry is received at Alpha is 0.60.
The probability that an enquiry is received at Beta is 0.25.
The probability that an enquiry is received at Gamma is 0.15.

Each enquiry is resolved at the station that receives the enquiry. The **percentages** of enquiries resolved within various times at **each** station are shown in the table.

		Time		
		$\leqslant$ 1 hour	**$\leqslant$ 24 hours**	**$\leqslant$ 72 hours**
Station	**Alpha**	55	80	100
	Beta	60	85	100
	Gamma	40	75	100

For example:
80 per cent of enquiries received at Alpha are resolved within 24 hours;
25 per cent of enquiries received at Alpha take between 1 hour and 24 hours to resolve.

(a) Find the probability that an enquiry, selected at random, is:

(i) resolved at Gamma; *(1 mark)*

(ii) resolved at Alpha within 1 hour; *(1 mark)*

(iii) resolved within 24 hours; *(2 marks)*

(iv) received at Beta, given that it is resolved within 24 hours. *(3 marks)*

(b) A random sample of 3 enquiries was selected.

Given that all 3 enquiries were resolved within 24 hours, calculate the probability that they were all received at:

(i) Beta; *(2 marks)*

(ii) the same station. *(4 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

4 The waiting time at a hospital's A&E department may be modelled by a normal distribution with mean μ and standard deviation $\frac{\mu}{2}$.

The department's manager wishes a 95% confidence interval for μ to be constructed such that it has a width of at most $0.2\,\mu$.

Calculate, to the nearest 10, an estimate of the minimum sample size necessary in order to achieve the manager's wish. *(5 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

5 An examination of 160 e-mails received by Gopal showed that 72 had attachments.

An examination of 250 e-mails received by Haley showed that 102 had attachments.

Stating **two** necessary assumptions about the selection of e-mails, construct an approximate 99% confidence interval for the difference between the proportion of e-mails received by Gopal that have attachments and the proportion of e-mails received by Haley that have attachments. *(8 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

6 The weight, X grams, of a dressed pheasant may be modelled by a normal random variable with a mean of 1000 and a standard deviation of 120.

Pairs of dressed pheasants are selected for packing into boxes. The total weight of a pair, $Y = X_1 + X_2$ grams, may be modelled by a normal distribution with a mean of 2000 and a standard deviation of 140.

(a) (i) Show that $\text{Cov}(X_1, X_2) = -4600$. *(3 marks)*

(ii) Given that $X_1 - X_2$ may be assumed to be normally distributed, determine the probability that the difference between the weights of a selected pair of dressed pheasants exceeds 250 grams. *(5 marks)*

(b) The weight of a box is independent of the total weight of a pair of dressed pheasants, and is normally distributed with a mean of 500 grams and a standard deviation of 40 grams.

Determine the probability that a box containing a pair of dressed pheasants weighs less than 2750 grams. *(5 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

7 (a) The random variable X has a Poisson distribution with $\mathrm{E}(X) = \lambda$.

(i) Prove, from first principles, that $\mathrm{E}(X(X-1)) = \lambda^2$. *(3 marks)*

(ii) Hence deduce that $\mathrm{Var}(X) = \mathrm{E}(X)$. *(2 marks)*

(b) The random variable Y has a Poisson distribution with $\mathrm{E}(Y) = 2.5$.

Given that $Z = 4Y + 30$:

(i) show that $\mathrm{Var}(Z) = \mathrm{E}(Z)$; *(3 marks)*

(ii) give a reason why the distribution of Z is **not** Poisson. *(1 mark)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

QUESTION PART REFERENCE

8 The tensile strength of rope is measured in kilograms.

The standard deviation of the tensile strength of a particular design of 10 mm diameter rope is known to be 285 kilograms. A retail organisation, which buys such rope from two manufacturers, A and B, wishes to compare their ropes for mean tensile strength.

The mean tensile strength, $\bar{x}$, of a random sample of 80 lengths from manufacturer A was 3770 kilograms.

The mean tensile strength, $\bar{y}$, of a random sample of 120 lengths from manufacturer B was 3695 kilograms.

(a) (i) Test, at the 5% level of significance, the hypothesis that there is no difference between the mean tensile strength of rope from manufacturer A and that of rope from manufacturer B. *(6 marks)*

(ii) Why was it **not** necessary to know the distributions of tensile strength in order for your test in part **(a)(i)** to be valid? *(1 mark)*

(b) (i) Deduce that, for your test in part **(a)(i)**, the critical values of $(\bar{x} - \bar{y})$ are ± 80.63, correct to two decimal places. *(2 marks)*

(ii) In fact, the mean tensile strength of rope from manufacturer A exceeds that of rope from manufacturer B by 125 kilograms.

Determine the probability of a Type II error for a test of the hypothesis in part **(a)(i)** at the 5% level of significance, based upon a random sample of 80 lengths from manufacturer A and a random sample of 120 lengths from manufacturer B. *(4 marks)*

QUESTION PART REFERENCE	
	.....
	.....
	.....
	.....
	.....
	.....
	.....
	.....
	.....

Do not write outside the box

QUESTION PART REFERENCE

QUESTION PART REFERENCE

END OF QUESTIONS

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	

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 1

Do not write outside the box

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**
Room	**Double**	0.15	0.35	0.45	0.05
	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

.....

.....

.....

.....

.....

Do not write outside the box

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

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

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 6

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 $\mathrm{E}(X) = np$. *(3 marks)*

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

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

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

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

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

Use a distributional approximation to estimate $\mathrm{P}(Y \leqslant k)$, 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

QUESTION PART REFERENCE

Answer space for question 7

END OF QUESTIONS

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	

General Certificate of Education
Advanced Level Examination
June 2013

Mathematics MS03

Unit Statistics 3

Tuesday 18 June 2013 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 number of telephone calls per hour to an out-of-hours doctors' service may be modelled by a Poisson distribution.

The total number of telephone calls received during a random sample of 12 weekday night shifts, all of the same duration, was 392 .

(a) Calculate an approximate 98% confidence interval for the mean number of calls received per weekday night shift. *(5 marks)*

(b) The mean number of calls received during weekend shifts of 48 hours' total duration is 136.8 .

Comment on a claim that the mean number of calls **per hour** during weekend shifts is greater than that during weekday night shifts, which are each of **14** hours' duration. *(3 marks)*

QUESTION PART REFERENCE

Answer space for question 1

QUESTION PART REFERENCE

Answer space for question 1

2 On a rail route between two stations, A and B, 90% of trains leave A on time and 10% of trains leave A late.

Of those trains that leave A on time, 15% arrive at B early, 75% arrive on time and 10% arrive late.

Of those trains that leave A late, 35% arrive at B on time and 65% arrive late.

(a) Represent this information by a fully-labelled tree diagram. *(3 marks)*

(b) Hence, or otherwise, calculate the probability that a train:

(i) arrives at B early or on time;

(ii) left A on time, given that it arrived at B on time;

(iii) left A late, given that it was not late in arriving at B. *(7 marks)*

(c) Two trains arrive late at B. Assuming that their journey times are independent, calculate the probability that exactly one train left A on time. *(4 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

QUESTION PART REFERENCE

Answer space for question 2

3 A builders' merchant's depot has two machines, X and Y, each of which can be used for filling bags with sand or gravel.

The weight, in kilograms, delivered by machine X may be modelled by a normal distribution with mean μ_X and standard deviation 25.

The weight, in kilograms, delivered by machine Y may be modelled by a normal distribution with mean μ_Y and standard deviation 30.

Fred, the depot's yardman, records the weights, in kilograms, of a random sample of 10 bags of sand delivered by machine X as

1055 1045 1000 985 1040 1025 1005 1030 1015 1060

He also records the weights, in kilograms, of a random sample of 8 bags of gravel delivered by machine Y as

1085 1055 1055 1000 1035 1050 1005 1075

(a) Construct a 95% confidence interval for $\mu_Y - \mu_X$, giving the limits to the nearest 5 kg. *(7 marks)*

(b) Dot, the depot's manager, commented that Fred's data collection may have been biased.

Justify her comment and explain how the possible bias could have been eliminated. *(2 marks)*

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

4 An analysis of a sample of 250 patients visiting a medical centre showed that 38 per cent were aged over 65 years.

An analysis of a sample of 100 patients visiting a dental practice showed that 21 per cent were aged over 65 years.

Assume that each of these two samples has been randomly selected.

Investigate, at the 5% level of significance, the hypothesis that the percentage of patients visiting the medical centre, who are aged over 65 years, exceeds that of patients visiting the dental practice, who are aged over 65 years, by more than 10 per cent. *(8 marks)*

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

5 The schedule for an organisation's afternoon meeting is as follows.

Session A (Speaker 1)	2.00 pm to 3.15 pm
Session B (Discussion)	3.15 pm to 3.45 pm
Session C (Speaker 2)	3.45 pm to 5.00 pm

Records show that:

the duration, X, of Session A has mean 68 minutes and standard deviation 10 minutes;

the duration, Y, of Session B has mean 25 minutes and standard deviation 5 minutes;

the duration, Z, of Session C has mean 73 minutes and standard deviation 15 minutes;

and that:

$$\rho_{XZ} = 0 \qquad \rho_{XY} = -0.8 \qquad \rho_{YZ} = 0$$

(a) Determine the means and the variances of:

(i) $L = X + Z$; *(2 marks)*

(ii) $M = X + Y$. *(3 marks)*

(b) Assuming that L and M are each normally distributed, determine the probability that:

(i) the total time for the two speaker sessions is less than $2\frac{1}{2}$ hours; *(2 marks)*

(ii) Session C is late in starting. *(3 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

6 The demand for a *WWSatNav* at a superstore may be modelled by a Poisson distribution with a mean of 2.5 per day. The superstore is open 6 days each week, from Monday morning to Saturday evening.

(a) (i) Determine the probability that the demand for *WWSatNav*s during a particular week is at most 18. *(2 marks)*

(ii) The superstore receives a delivery of *WWSatNav*s on each Sunday evening. The manager, Meena, requires that the probability of *WWSatNav*s being out of stock during a week should be at most 5%.

Determine the minimum number of *WWSatNav*s that Meena requires to be in stock after a delivery. *(2 marks)*

(b) (i) Use a distributional approximation to estimate the probability that the demand for *WWSatNav*s during a period of **2 weeks** is more than 35. *(4 marks)*

(ii) Changes to the superstore's delivery schedule result in it receiving a delivery of *WWSatNav*s on alternate Sunday evenings. Meena now requires that the probability of *WWSatNav*s being out of stock during the 2 weeks following a delivery should be at most 5%.

Use a distributional approximation to determine the minimum number of *WWSatNav*s that Meena now requires to be in stock after a delivery. *(3 marks)*

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

7 It is claimed that the proportion, P, of people who prefer cooked fresh garden peas to cooked frozen garden peas is greater than 0.50.

(a) In an attempt to investigate this claim, a sample of 50 people were each given an unlabelled portion of cooked fresh garden peas and an unlabelled portion of cooked frozen garden peas to taste. After tasting each portion, the people were each asked to state which of the two portions they preferred.

Of the 50 people sampled, 29 preferred the cooked fresh garden peas.

Assuming that the 50 people may be considered to constitute a random sample, use a binomial distribution and the 10% level of significance to investigate the claim. *(6 marks)*

(b) It was then decided to repeat the tasting in part **(a)** but to involve a sample of 500, rather than 50, people.

Of the 500 people sampled, 271 preferred the cooked fresh garden peas.

(i) Assuming that the 500 people may be considered to constitute a random sample, use an approximation to the distribution of the sample proportion, $\widehat{P}$, and the 10% level of significance to again investigate the claim. *(4 marks)*

(ii) The critical value of $\widehat{P}$ for the test in part **(b)(i)** is 0.529, correct to three significant figures. It is also given that, in fact, 55 per cent of people prefer cooked fresh garden peas.

Estimate the power for a test of the claim that $P > 0.50$ based on a random sample of 500 people and using the 10% level of significance. *(5 marks)*

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

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 $\mathbf{96\%}$ 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

Do not write outside the box

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

Do not write outside the box

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]

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 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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

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

Do not write outside the box

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

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

QUESTION PART REFERENCE

Answer space for question 6

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

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

END OF QUESTIONS

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

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

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 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

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

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

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

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

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

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

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

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

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

END OF QUESTIONS

Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname _____

Forename(s) _____

Candidate signature _____

A-level MATHEMATICS

Unit Statistics 3

Monday 27 June 2016 Morning Time allowed: 1 hour 30 minutes

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

You may use a graphics calculator.

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 In advance of a referendum on independence, the regional assembly of an eastern province of a particular country carried out an opinion poll to assess the strength of the 'Yes' vote.

Of the 480 men polled, 264 indicated that they intended to vote 'Yes', and of the 500 women polled, 220 indicated that they intended to vote 'Yes'.

(a) Construct an approximate 95% confidence interval for the difference between the proportion of men who intend to vote 'Yes' and the proportion of women who intend to vote 'Yes'.

[6 marks]

(b) Comment on a claim that, in the forthcoming referendum, the percentage of men voting 'Yes' will exceed the percentage of women voting 'Yes' by at least 2.5 per cent. Justify your answer.

[2 marks]

QUESTION PART REFERENCE	**Answer space for question 1**

Do not write outside the box

QUESTION PART REFERENCE	**Answer space for question 1**

2 A plane flies regularly between airports D and T with an intermediate stop at airport M. The time of the plane's departure from, or arrival at, each airport is classified as either early, on time, or late.

On 90% of flights, the plane departs from D on time, and on 10% of flights, it departs from D late.

Of those flights that depart from D on time, 65% then depart from M on time and 35% depart from M late.

Of those flights that depart from D late, 15% then depart from M on time and 85% depart from M late.

Any flight that departs from M on time has probability 0.25 of arriving at T early, probability 0.60 of arriving at T on time and probability 0.15 of arriving at T late.

Any flight that departs from M late has probability 0.10 of arriving at T early, probability 0.20 of arriving at T on time and probability 0.70 of arriving at T late.

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

[3 marks]

(b) Hence, or otherwise, calculate the probability that the plane:

(i) arrives at T on time;

(ii) arrives at T on time, given that it departed from D on time;

(iii) does not arrive at T late, given that it departed from D on time;

(iv) does not arrive at T late, given that it departed from M on time.

[8 marks]

(c) Three independent flights of the plane depart from D on time.

Calculate the probability that two flights arrive at T on time and that one flight arrives at T early.

[4 marks]

QUESTION PART REFERENCE	**Answer space for question 2**

Do not write outside the box

QUESTION PART REFERENCE	**Answer space for question 2**

QUESTION PART REFERENCE	**Answer space for question 2**

QUESTION PART REFERENCE

Answer space for question 2

3 Car parking in a market town's high street was, until 31 May 2014, limited to one hour free of charge between $8\,\text{am}$ and $6\,\text{pm}$. Records show that, during a period of 30 days prior to this date, a total of 315 penalty tickets were issued.

Car parking in the high street later became limited to thirty minutes free of charge between $8\,\text{am}$ and $6\,\text{pm}$. A subsequent investigation revealed that, during a period of 60 days from 1 October 2014, a total of 747 penalty tickets were issued.

The daily numbers of penalty tickets issued may be modelled by independent Poisson distributions with means λ_A until 31 May 2014 and λ_B from 1 October 2014.

Investigate, at the 1% level of significance, a claim by traders on the high street that $\lambda_B > \lambda_A$.

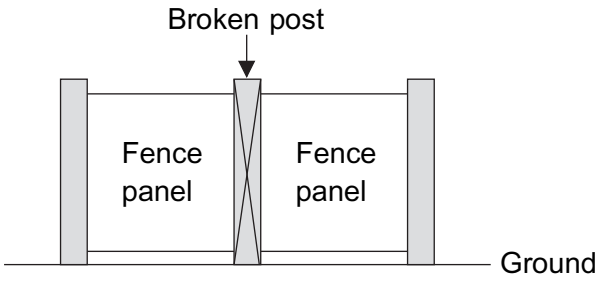
[7 marks]

QUESTION PART REFERENCE

Answer space for question 3

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

4 Ben is a fencing contractor who is often required to repair a garden fence by replacing a broken post between fence panels, as illustrated.



The tasks involved are as follows.

U : detach the two fence panels from the broken post
V : remove the broken post
W : insert a new post
X : attach the two fence panels to the new post

The mean and the standard deviation of the time, in minutes, for each of these tasks are shown in the table.

Task	Mean	Standard deviation
U	15	5
V	40	15
W	75	20
X	20	10

The random variables U, V, W and X are pairwise independent, except for V and W for which $\rho_{VW} = 0.25$.

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

(i) $R = U + X$;

(ii) $F = V + W$;

(iii) $T = R + F$;

(iv) $D = W - V$.

[8 marks]

(b) Assuming that each of R, F, T and D is approximately normally distributed, determine the probability that:

(i) the total time taken by Ben to repair a garden fence is less than 3 hours;

(ii) the time taken by Ben to insert a new post is at least 1 hour more than the time taken by him to remove the broken post.

[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

5 (a) The random variable X, which has distribution $\mathrm{N}(\mu_X, \sigma^2)$, is independent of the random variable Y, which has distribution $\mathrm{N}(\mu_Y, \sigma^2)$.

In order to test $\mathrm{H}_0: \mu_X = 1.5\mu_Y$, samples of size n are taken on each of X and Y and the random variable $\overline{D}$ is defined as

$$\overline{D} = \overline{X} - 1.5\overline{Y}$$

State the distribution of $\overline{D}$ assuming that H_0 is true.

[4 marks]

(b) A machine that fills bags with rice delivers weights that are normally distributed with a standard deviation of 4.5 grams.

The machine fills two sizes of bags: large and extra-large.

The mean weight of rice in a random sample of 50 large bags is 1509 grams.

The mean weight of rice in an independent random sample of 50 extra-large bags is 2261 grams.

Test, at the 5% level of significance, the claim that, on average, the rice in an extra-large bag is $1\frac{1}{2}$ times as heavy as that in a large bag.

[6 marks]

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

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

6 (a) The discrete random variable X has probability distribution given by

$$\mathrm{P}(X = x) = \begin{cases} \dfrac{\mathrm{e}^{-\lambda}\lambda^{x}}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Show that $\mathrm{E}(X) = \lambda$ and that $\mathrm{Var}(X) = \lambda$.

[7 marks]

(b) In light-weight chain, faults occur randomly and independently, and at a constant average rate of 0.075 per metre.

(i) Calculate the probability that there are no faults in a 10-metre length of this chain.

[2 marks]

(ii) Use a distributional approximation to estimate the probability that, in a 500-metre reel of light-weight chain, there are:

(A) fewer than 30 faults;

(B) at least 35 faults but at most 45 faults.

[7 marks]

(c) As part of an investigation into the quality of a new design of medium-weight chain, a sample of **fifty** 10-metre lengths was selected.

Subsequent analysis revealed a total of 49 faults.

Assuming that faults occur randomly and independently, and at a constant average rate, construct an approximate 98% confidence interval for the average number of faults **per metre**.

[6 marks]

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

QUESTION PART REFERENCE

Answer space for question 6

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

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

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED