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 2014

Statistics SS04

Unit Statistics 4

Monday 16 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 National records show that 35 per cent of train passengers buy their tickets in advance. A random sample of 25 passengers using a particular railway station is selected, and it is found that 13 of them bought their tickets in advance.

(a) Investigate, at the 10% level of significance, whether the data support the view that the percentage of passengers from this station who buy their tickets in advance is different from the national figure of 35 per cent.

[6 marks]

(b) It was suggested that, for a follow-up survey, it would be easier to collect data from all the passengers in a particular railway carriage. Explain in context why it would **not** be appropriate to then apply the test that you have used in part **(a)**.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 1

QUESTION PART REFERENCE

Answer space for question 1

2 A company manufactures blank compact discs for the recording industry. The quality of these discs is such that 3 per cent have small surface scratches on the recording side and 25 per cent have small surface scratches on the printing side.

(a) A small recording studio orders 200 discs, which may be regarded as a random sample from the population of all discs produced by the company.

Use an approximate distribution to estimate the probability that more than 4 of these 200 discs will have scratches on the **recording** side.

[4 marks]

(b) A larger recording studio orders $10\,000$ discs, which may also be regarded as a random sample from the population of all discs produced by the company.

Use an approximate distribution to estimate the probability that at most 2550 of these $10\,000$ discs will have scratches on the **printing** side.

[6 marks]

QUESTION PART REFERENCE

Answer space for question 2

QUESTION PART REFERENCE

Answer space for question 2

3 Kelly makes birthday cards which she sells on her website. Assume that the daily number of cards that Kelly sells can be modelled by a Poisson distribution with mean λ.

(a) (i) On a randomly chosen day, Kelly sells 58 cards. Construct an approximate 95% confidence interval for λ.

[3 marks]

(ii) Using this interval, comment on Kelly's claim that she sells an average of 75 cards per day.

[2 marks]

(b) Would your answer to part **(a)(ii)** change if it were based instead on an approximate 90% confidence interval for λ? You are **not** required to construct another interval but you should justify your answer.

[2 marks]

(c) Give **two** reasons why the confidence interval constructed in part **(a)(i)** is approximate rather than exact.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

4 During weekday mornings, customers arrive at Hanif's newsagents at random at a constant average rate of 1 customer every 3 minutes. The amount spent per visit by each customer can be modelled by a normal distribution with mean £4.89 and standard deviation £0.92. The amounts spent are independent from customer to customer.

(a) (i) Find the probability that, during a 30-minute period on a weekday morning, exactly 10 customers will arrive at Hanif's newsagents.

[3 marks]

(ii) Find the probability that the **total** amount spent by 10 customers will be less than £45.

[6 marks]

(iii) What is the probability that, between 10.45 am and 11.15 am on a given Wednesday, exactly 10 customers will arrive **and** a total of **at least** £45 will be spent by these 10 customers?

[2 marks]

(b) After a sales promotion, Hanif believes that the mean amount spent is now more than £5. To investigate this, he records the amounts spent by a random sample of 7 customers. These amounts (in £) are as follows.

5.42 2.50 4.98 6.72 5.09 7.15 6.14

Investigate Hanif's belief at the 5% level of significance.

Assume that the amount spent per visit by each customer can be modelled by a normal distribution with unknown standard deviation.

[8 marks]

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

5 A magazine reported: "*Three quarters of the world population now have access to a mobile phone*."

(a) A survey of 350 people in the UK showed that 78 per cent of them had access to a mobile phone.

Use a distributional approximation and the 5% level of significance to investigate whether the percentage in the UK is greater than the percentage reported for the world as a whole.

[8 marks]

(b) A survey of 125 people in a particular developing country showed that 55 of them had access to a mobile phone.

(i) Construct an approximate 95% confidence interval for p, the proportion in the population of this developing country with access to a mobile phone.

[5 marks]

(ii) Does the interval calculated in part **(b)(i)** support a claim that more than a third of the population of the developing country has access to a mobile phone? Justify your answer.

[2 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

QUESTION PART REFERENCE

Answer space for question 5

6 (a) Jennifer uses two different methods to mow her lawn.

Method A is a two-stage process involving cutting the grass with a strimmer then collecting the grass by raking it up. The time taken to cut the grass is a normally distributed random variable, U, with mean 40 minutes and standard deviation 4.2 minutes. Once cut, the time taken to collect the grass is a normally distributed random variable, V, with mean 30 minutes and standard deviation 2.8 minutes. Thus the total time Jennifer takes to mow her lawn using Method A is given by $U + V$.

Method B uses a mower with a rechargeable battery that will cut and collect the grass at the same time. The time taken to do this is a normally distributed random variable, W, with mean 48 minutes and standard deviation 2.2 minutes. In addition to this, the battery has to be recharged once during the cut, and this time is fixed at 10 minutes. Thus the total time Jennifer takes to mow her lawn using Method B is given by $W + 10$.

The random variable X is defined by $X = U + V - W - 10$.

Assume that U, V and W are independent from each other.

(i) Find the mean and the variance of X.

[3 marks]

(ii) Hence find $\mathrm{P}(X > 5)$.

[3 marks]

(iii) Explain in the context of the question what your answer in part **(a)(ii)** represents.

[2 marks]

(b) Jennifer decides to try out **Method C**, which uses a petrol lawn mower. This will also cut and collect the grass at the same time but does not require a battery recharge. The time taken for Method C is a normally distributed random variable, Y, with mean 45 minutes and standard deviation 3.3 minutes.

The petrol mower can be assumed to use exactly 2 litres of fuel per hour.

Find the probability that, on a particular occasion when Jennifer mows her lawn using Method C, the mower uses between 1.5 and 1.75 litres of fuel.

[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