AQA

Please write clearly in block capitals.

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

Surname _____

Forename(s) _____

Candidate signature _____

A-level STATISTICS

Unit Statistics 5

Friday 24 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 Jamal supervises a production line in a factory that produces individual tubs of ice cream.

During a hot spell of weather when the demand for tubs of ice cream has been especially high, Jamal suspects that the amounts dispensed into tubs are more variable than usual.

In order to check this, he takes a random sample of 14 tubs and records the volume, in millilitres, of ice cream in each tub. The results are as follows.

117.8	126.3	119.4	116.7	124.7	120.1	118.5
125.9	123.8	121.3	127.6	117.9	126.8	125.9

(a) (i) Construct a 98% confidence interval for the current variance of the volume of ice cream in a tub.

You may assume that the volumes follow a normal distribution.

[6 marks]

(ii) Long term records reveal that the variance of the volume of ice cream in a tub was previously $5\,\text{ml}^2$.

Comment, with justification, on what Jamal can conclude about his suspicion.

[2 marks]

(b) Jamal's colleague, Kajika, constructs a **95%** confidence interval for the variance of the volume of ice cream in a tub based on the above 14 results.

Without further calculation, state whether Kajika's confidence interval will lead her to the same conclusion as Jamal. Give a reason for your answer.

[2 marks]

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

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

2 Matthew is a primary school teacher. Every year, as part of his teaching of units of length, he asks each of his class of Year 5 children to place in the playground, independently and without guidance, two markers at a distance apart that is estimated by the child to be one metre.

Experience, over many years, has led Matthew to conclude that the estimated distance, X cm, can be modelled by the probability density function

$$f(x) = \begin{cases} k & 80 \leqslant x \leqslant 130 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(a) (i) Show that the exact value of k is $\frac{1}{50}$.

(ii) State the probability that $X = 100$. Interpret your answer in the context of the question.

[4 marks]

(b) Find the mean and the standard deviation of X.

[2 marks]

(c) Matthew measures the value x, for each child, and records the child's score as $(x - 100)$ when $x \geqslant 100$, but as $(100 - x)$ when $x < 100$.

He awards a gold star to any child whose score is less than 5.

(i) Find the probability that a child is awarded a gold star.

(ii) In a year when Matthew's class has 25 children, find the probability that he awards 6 or more gold stars.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 2

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

3 Lucy practises her golf swing at the local driving range. She has the choice of using two brands of ball, *Whizzer and Screamer*. She believes that her driving distance when she uses a *Whizzer* ball is **more variable** than her driving distance when she uses a *Screamer* ball.

For each brand of ball, Lucy records a sample of her driving distances, in yards, as follows.

Whizzer	164	171	212	191	186	189	211	170
	208	157	156	205	201	197	154	140
Screamer	191	192	184	175	185	163	199	171
	189	155	172	186				

(a) Investigate Lucy's belief using the 5% significance level.

[8 marks]

(b) State the assumptions that are necessary in order for the test in part **(a)** to be valid.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

4 Helena is the manager of a shopping mall on the outskirts of a large town. She is interested in the distances that her customers travel to shop in the mall. She surveys a random sample of 200 customers and records their distances travelled, d kilometres. Her results are summarised in **Table 1**.

Table 1

d	Frequency
0–1	8
2–3	13
4–6	27
7–9	37
10–14	39
15–19	45
20–24	23
25–29	5
30–49	3

Helena decided to use a χ^2-test to determine whether a normal distribution would provide a suitable model for the distances travelled by customers.

From the above data, she calculated correctly an estimate of the mean to be $12.31\,\text{km}$ and an estimate of the standard deviation to be $7.40\,\text{km}$.

Using these values and a normal distribution, she then calculated the expected frequencies, each correct to two decimal places, given in **Table 2**.

Table 2

d	Expected frequency
$\leqslant 1.5$	w
2–3	8.98
4–6	19.85
7–9	y
10–14	52.86
15–19	43.60
20–24	23.17
25–29	7.93
> 29.5	z

(a) (i) Calculate values for the missing frequencies w, y and z in **Table 2**.

[6 marks]

(ii) Carry out Helena's χ^2-test at the 5% significance level.

[8 marks]

(b) Helena's publicity manager, Mike, subsequently ran an advertising campaign for the shopping mall.

Following the campaign, he surveyed another random sample of 200 customers and recorded their distances travelled. He carried out a hypothesis test to investigate whether the mean of the distances travelled by customers had changed as a result of his advertising campaign.

Mike claimed that, for his test to be valid, it was not necessary for the distances travelled to follow a normal distribution.

State, with a reason, whether his claim is correct.

[2 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**

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

5 The time, T seconds, between the arrival of successive vehicles at a zebra crossing on a road through a village can be modelled by an exponential distribution with parameter $\lambda = 0.025$.

(a) Write down the mean and the variance of T.

[2 marks]

(b) Geoffrey is an elderly pedestrian who takes 30 seconds to cross the road using this zebra crossing.

(i) Calculate the probability that no vehicle arrives whilst Geoffrey is crossing.

(ii) Calculate the probability that no vehicle arrives whilst Geoffrey makes two independent crossings.

[3 marks]

(c) Tandeep, a local resident, monitors the traffic through the village. He records the times between 75 successive vehicles arriving at the zebra crossing.

Use a distributional approximation to estimate the probability that the **mean** of the times recorded by Tandeep exceeds 35 seconds.

[4 marks]

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

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

6 (a) Lian is a student at a sixth form college. As part of a project, she is investigating the Body Mass Index (BMI) of female students at the college.

Lian records the BMI, in $\mathrm{kg/m^2}$, of the 15 female students in her class with the following results.

22.3	25.2	23.7	22.0	24.1	22.5	25.3	24.8
23.1	23.5	25.6	25.8	25.2	25.4	24.3	

The BMI of female sixth form college students may be assumed to be normally distributed with a standard deviation of $4\,\mathrm{kg/m^2}$. Lian wishes to test whether her data comes from a population with this standard deviation.

(i) Carry out a hypothesis test for Lian at the 1% level of significance.

[8 marks]

(ii) Give a reason, based upon Lian's collection of data, why your conclusion in part **(a)(i)** may not be reliable.

[1 mark]

QUESTION PART REFERENCE

Answer space for question 6(a)

QUESTION PART REFERENCE	**Answer space for question 6(a)**

Question 6 continues on the next page

(b) Stephen, another student at the college, read an article in a newspaper stating that male adults who travel to work by car have, on average, a BMI more than $1\,\text{kg/m}^2$ greater than those who travel to work by alternative means of transport.

To investigate whether this statement also applies to male students travelling to the college, Stephen calculated the BMI of each student in two independent random samples of male students. The 11 students in one sample travelled to the college by car, whereas the 8 students in the other sample travelled to the college by alternative means of transport. His results are given in the table.

BMI of those travelling by car	28.4 23.1	26.7 26.7	25.2 24.1	27.8 27.6	22.8 25.8	28.9
BMI of those travelling by alternative means	26.7 22.6	23.9 23.1	23.0	22.5	25.7	26.3

Making any necessary assumptions, investigate, at the 5% significance level, the hypothesis that male students who travel to the college by car have a BMI which is, on average, more than $1\,\text{kg/m}^2$ greater than those who travel to the college by alternative means of transport.

[11 marks]

QUESTION PART REFERENCE

Answer space for question 6(b)

QUESTION PART REFERENCE

Answer space for question 6(b)

QUESTION PART REFERENCE	**Answer space for question 6(b)**

END OF QUESTIONS