AQA

Please write clearly in block capitals.

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

Surname _____

Forename(s) _____

Candidate signature _____

A-level MATHEMATICS

Unit Statistics 4

Wednesday 29 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 The random variable X has an exponential distribution with mean 16 .

Find:

(a) $\mathrm{P}(X < 10)$;

(b) $\mathrm{P}(10 < X < 20)$;

(c) $\mathrm{P}(X \neq 15)$.

[5 marks]

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

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

2 Jian is a medical graduate who is observing the work of doctors at a surgery. He is interested in factors affecting patients' blood pressures and suspects that a patient's diastolic blood pressure may be affected by whether the patient is sitting or is standing.

Jian selects a sample of 12 patients from those attending the surgery's blood pressure clinic. He records the diastolic blood pressure of each of these patients when sitting and then again after standing for two minutes. The results, in mm Hg, are shown in the table.

Patient	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
Sitting	72	79	69	76	82	66	87	78	85	77	84	70
Standing	81	70	79	84	88	77	79	85	91	74	88	80

(a) Construct a 95% confidence interval for the mean difference between a patient's diastolic blood pressure when sitting and when standing. Assume that the differences in blood pressure are normally distributed.

[8 marks]

(b) State, with a reason, whether your confidence interval supports Jian's suspicion.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 2

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

3 A manufacturer of doorknobs uses a process that is designed to produce a doorknob with a diameter of $6\,\text{cm}$. Measurements over time have shown that the standard deviation of the diameter is $0.10\,\text{cm}$.

In an effort to reduce the variation in diameter, a redesigned process was implemented. The diameter, x centimetres, of each of a random sample of 25 doorknobs produced by the redesigned process was measured, and these values gave the following summarised information.

$$\sum (x - \bar{x})^2 = 0.071306$$

(a) Stating a necessary assumption, construct a 98% confidence interval for the standard deviation of the diameter of doorknobs produced by the redesigned process.

[7 marks]

(b) What does your confidence interval reveal? Justify your answer.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

4 Maureen is the owner of a small independent garage called *ThirCars*. She believes that cars serviced at her garage are, on average, more than 3 years older than cars serviced at the national chain *NYAutos*.

To investigate her belief, Maureen records the ages of a random sample of 12 cars that have recently been serviced at her garage. She also obtains the ages of a random sample of 12 cars that have recently been serviced at *NYAutos*. The results, **in months**, are given in the table.

ThirCars	65	60	90	72	81	98	77	78	107	75	82	75
NYAutos	10	35	42	32	26	33	53	55	18	29	33	42

The two samples of car ages may be modelled by independent normal distributions, *ThirCars* with variance σ_T^2 and *NYAutos* with variance σ_N^2.

(a) Show, using the 5% level of significance, that it may be assumed that $\sigma_T^2 = \sigma_N^2$.

[8 marks]

(b) Hence investigate, at the 5% level of significance, Maureen's belief that cars serviced at *ThirCars* are, on average, more than 3 years older than cars serviced at *NYAutos*.

[11 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 The random variable X has a geometric distribution with parameter p. The mean and variance of X are denoted by μ and σ^2 respectively.

(a) (i) Prove that $\mu = \frac{1}{p}$.

(ii) Prove that $\mathrm{E}(X(X-1)) = \frac{2(1-p)}{p^2}$, and hence deduce an expression for σ^2.

[7 marks]

(b) Given that $p = 0.1$, calculate $\mathrm{P}(\mu - \sigma < X < \mu + \sigma)$.

[4 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 The random variable R denotes the outcome of a trial. It is claimed that R can be modelled by the probability distribution

$$\mathrm{P}(R = r) = \begin{cases} \dfrac{\binom{6}{r}\binom{6}{4-r}}{495} & r = 0, 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

To investigate this claim, the value of R was recorded for a sample of 99 independent trials, with the following results.

r	0	1	2	3	4
Frequency	2	18	41	32	6

Use a χ^2 goodness of fit test and the 10% level of significance to test the claim that R can be modelled by the probability distribution given above.

[12 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 The random variable X denotes the number of successes in a random sample of size n from a large population in which the proportion of successes is p.

The random variable Y denotes the number of successes in an independent random sample of size $3n$ from the same large population.

(a) Write down expressions for the mean and the variance of **each** of X and Y.

[2 marks]

(b) Two estimators suggested for p are

$$\widehat{P}_1 = \frac{X+Y}{4n} \quad \text{and} \quad \widehat{P}_2 = \frac{1}{2}\left(\frac{X}{n} + \frac{Y}{3n}\right)$$

Show that both $\widehat{P}_1$ and $\widehat{P}_2$ are unbiased and consistent estimators for p, and that $\widehat{P}_1$ is the more efficient.

[7 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

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