OXFORD
INTERNATIONAL
AQA EXAMINATIONS

Please write clearly in block capitals.

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

Surname ____________________

Forename(s) ____________________

Candidate signature ____________________
I declare this is my own work.

INTERNATIONAL A-LEVEL
FURTHER MATHEMATICS

(9665/FM04) Unit FS2 Statistics

Monday 12 June 2023 07:00 GMT Time allowed: 1 hour 30 minutes

Materials
- For this paper you must have the Oxford International AQA Booklet of Formulae and Statistical Tables (enclosed).
- You may use a graphical 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.
- You must answer the questions in the spaces provided. Do not write outside the box around each page or on blank pages.
- If you need extra space for your answer(s), use the lined pages at the end of this book. Write the question number against your answer(s).
- Do all rough work in this book. Cross through any work you do not want to be marked.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 80.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- Show all necessary working; otherwise marks may be lost.

For Examiner's Use	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
8	
9	
TOTAL	

There are no questions on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

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

1 The masses of apples from an orchard are assumed to be normally distributed with mean μ grams and standard deviation σ grams.

A random sample of 10 apples is taken and their masses, in grams, $M_i \{i = 1, 2, ..., 10\}$ measured.

You may assume the M_i variables are independent.

1 (a) Complete the table below of sampling distributions for the statistics shown.

[3 marks]

Statistic	M_1	$M_{10} - M_1$	$\sum_{i=1}^{10} M_i$	$\frac{1}{10}\sum_{i=1}^{10} M_i$
Sampling Distribution	$\mathrm{N}(\mu, \sigma^2)$			

1 (b) It is given that $X_i = \dfrac{M_i - \mu}{\sigma}$

1 (b) (i) Write down the distribution of X_i

[1 mark]

Answer ____________________

1 (b) (ii) Give a reason why X_i is not a statistic.

[1 mark]

5

Turn over ►

2 An engineering company is seeking to purchase a new machine.

The company will choose between two machines: machine *A* or machine *B*

The company will choose the machine which has the smallest variance in the length, x micrometres, of components manufactured on the machine.

Random samples of components manufactured by machine *A* and machine *B* are taken.

Assume that the populations from which the random samples are taken are normal and independent distributions.

Summary results of the samples are shown in the table below.

	Sample size	$\sum(x-\bar{x})^2$
Machine *A*	11	556.3
Machine *B*	7	341.4

Test at the 5% level of significance whether there is evidence for the company to select one machine rather than the other.

[8 marks]

Do not write outside the box

8

Turn over ►

Do not write outside the box

3 The random variable X has a Poisson distribution $\text{Po}(\lambda)$

A test is conducted at the 5% level of significance with the hypotheses

$$H_0 : \lambda = 1.8$$
$$H_1 : \lambda > 1.8$$

3 (a) Verify that the critical region is $X \geq 5$

[2 marks]

Do not write outside the box

3 (b) The actual value of λ is 3.4

Find the probability that a Type II error is made, giving your answer to three significant figures.

[2 marks]

Answer ____________________

3 (c) Find the power of the test, giving your answer to three significant figures.

[1 mark]

Answer ____________________

5

Turn over for the next question

4 Two random variables X and Y are independent with

$$X \sim \mathrm{N}\left(\mu_X, 5^2\right) \quad \text{and} \quad Y \sim \mathrm{N}\left(\mu_Y, 7^2\right)$$

A random sample of size 50 is taken from X and a random sample of size 14 is taken from Y

The sample means $\bar{X}$ and $\bar{Y}$ are calculated for both samples and the difference of the means, $\bar{X} - \bar{Y}$ is calculated.

4 (a) Write down $\mathrm{E}\left(\bar{X} - \bar{Y}\right)$ in terms of μ_X and μ_Y

[1 mark]

Answer ____________________

4 (b) Show that $\mathrm{Var}\left(\bar{X} - \bar{Y}\right) = 4$

[2 marks]

4 (c) Explain why $\bar{X} - \bar{Y}$ is normally distributed.

[1 mark]

Do not write outside the box

4 (d) For one particular sample of size 50 from X the value of $\overline{X}$ is 24.3

For one particular sample of size 14 from Y the value of $\overline{Y}$ is 28.5

Test at the 6% significance level whether X and Y have different population means.

[6 marks]

10

Turn over ►

5 A new variety of tomato plant has been developed to try to increase the yield of tomato fruit compared with the original variety.

A trial is conducted which takes a random sample of plants of the original variety and a random sample of the new variety. You may assume that the samples are from normal populations.

The table below shows the summary statistics for the samples.

Variety	**Number of plants**	**Sample mean yield (kilograms)**	**Sample variance s^2 (kilograms^2)**
Original	9	2.6	0.09
New	5	3.1	0.12

Investigate at the 1% level of significance whether the yield of tomato fruit is greater for the new variety compared with the original variety.

[9 marks]

Do not write outside the box

9

Turn over ►

Do not write outside the box

6 A random sample of size n is taken from a population.

Another random sample of size n is taken from a second independent population.

The first population has mean μ and variance σ^2
The sample mean from this population is $\overline{X}$

The second population has mean $a\mu$ and variance $b\sigma^2$ where a and b are positive real numbers. The sample mean from this population is $\overline{Y}$

6 (a) The statistic $S = 3\overline{X} - 4\overline{Y}$ is an unbiased estimator of μ

Find the value of a

[3 marks]

Answer ____________________

6 (b) Show that the statistic S is also a consistent estimator of μ

[3 marks]

Do not write outside the box

6 (c) The statistic $T = \frac{3}{2}\overline{X} - \overline{Y}$ is an unbiased estimator of μ

6 (c) (i) Show that the efficiency of T relative to S is $\frac{36 + 64b}{9 + 4b}$

[2 marks]

6 (c) (ii) Prove that the efficiency of T relative to S is greater than 4 and less than 16

[3 marks]

6 (c) (iii) State with a reason which of S or T is the more efficient estimator of μ

[1 mark]

12

Do not write outside the box

7 An independent agency is investigating published CO_2 emission values for a particular car model.

A random sample of 11 cars of the model are investigated and the CO_2 emission value, X grams per km, is measured.

Summary statistics from the investigation are

$$\sum x = 1034 \quad \text{and} \quad \sum x^2 = 97\,846$$

7 (a) Construct a 95% confidence interval for the mean CO_2 emission value of this car model.

Give your values to one decimal place.

[6 marks]

Answer ______________________________

Do not write outside the box

7 (b) State an assumption you have made about the distribution of CO_2 emission values for your confidence interval calculation from **part (a)** to be valid.

[1 mark]

7 (c) All car models with a mean CO_2 emission value of 100 grams per km or below qualify for a partial government refund.

7 (c) (i) State with a reason if your confidence interval provides evidence that the car model being investigated should qualify for the refund.

[2 marks]

7 (c) (ii) The car model is now investigated further by increasing the sample size to 65 cars.

The new 95% confidence interval is (97.3, 101.1)

Explain how this new confidence interval affects your answer to **part (c)(i)**.

[1 mark]

10

Turn over ►

8 A researcher measures the distance between mutations on a DNA strand.

They believe that the distances follow an exponential distribution with $\lambda = \dfrac{1}{1500}$ micrometres^{-1} and use this to determine expected distribution frequencies.

The measured and expected distribution frequencies from 1394 mutations are shown in the table below.

Interval (micrometres)	$<$ **500**	**500 to 1000**	**1000 to 1500**	**1500 to 2000**	**2000 to 2500**	**2500 to 3000**	$>$ **3000**
Measured frequency (M)	407	309	195	125	109	70	179
Expected frequency (E)	395	283	203	145	104	75	189
$(M-E)^2$	144	676	64	400	25	25	100

8 (a) Use integration to verify that the expected frequency for the 1500 to 2000 micrometres interval is 145 to the closest integer.

[2 marks]

8 (b) Determine the χ^2 test statistic using the values from the table for $(M-E)^2$

Give your answer correct to three significant figures.

[2 marks]

Answer ______________________________

Do not write outside the box

8 (c) Investigate the researcher's belief that the distances follow an exponential distribution with $\lambda = \frac{1}{1500}$ micrometres^{-1} at the 10% level of significance.

[5 marks]

9

Turn over for the next question

Do not write outside the box

9 A discrete random variable X has a Poisson distribution with population parameter λ

Its moment generating function is given by

$$\mathrm{M}_X(t) = \mathrm{e}^{\lambda\left(\mathrm{e}^t - 1\right)}$$

9 (a) (i) Use $\mathrm{M}_X(t)$ to show that $\mathrm{E}(X) = \lambda$

[2 marks]

9 (a) (ii) Use $\mathrm{M}_X(t)$ to determine $\mathrm{Var}(X)$

[3 marks]

Answer ____________________

9 (b) A second discrete random variable Y has a Poisson distribution with population parameter μ

The random variables X and Y are independent.

9 (b) (i) Find an expression for the moment generating function $\mathrm{M}_Z(t)$ of the random variable $Z = X + Y$

[2 marks]

Answer

9 (b) (ii) Using your answer for **part (b)(i)** verify that $Z \sim \mathrm{Po}(v)$ writing down v in terms of λ and μ

[1 mark]

Answer

Question 9 continues on the next page

9 (c) In this part you may use, without proof, the following results:

- $\lim_{n\to\infty}\left(1+\frac{\alpha}{n}\right)^n = e^{\alpha}$ where α is a constant
- If a random variable has the moment generating function of a Poisson random variable with parameter λ then it also has a Poisson probability distribution with parameter λ

A discrete random variable W has a binomial distribution with parameters n and p

Its moment generating function is $M_W(t) = \left(1+p\left(e^t - 1\right)\right)^n$

Prove that as $n \to \infty$ the probability distribution of W tends to the Poisson distribution with parameter $\lambda = np$ given that both λ and t are constant during the limiting process.

[3 marks]

9 (d) Interpret your result from **part (c)**.

[1 mark]

12

END OF QUESTIONS

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

Question number	**Additional page, if required. Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required. Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**