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 17 January 2022 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.

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

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.

Do not write outside the box

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

1 A random sample X is taken from a population with an unknown distribution.

Another random sample Y is taken from a second population, again with an unknown distribution.

The table below shows details of the two random samples.

Sample	Sample Size	Sample Mean	Sample Variance
X	100	9.5	4.2
Y	120	10.1	4.8

The samples are used to conduct a test at the 1% level of significance with the hypotheses

$$H_0: \quad \mu_Y = \mu_X$$
$$H_1: \quad \mu_Y > \mu_X$$

where μ_X is the population mean of sample X

and μ_Y is the population mean of sample Y

1 (a) Determine the critical value for the test, giving your answer to four significant figures.

[1 mark]

Do not write outside the box

1 (b) Determine the conclusion of the test.

[4 marks]

1 (c) State an assumption you have made to conduct the test.

[1 mark]

6

Turn over for the next question

Do not write outside the box

2 An international football tournament is taking place in a particular (host) country.

A total of 500 randomly selected football supporters were asked to predict which team will win the tournament.

Of these 500 supporters, 300 were from the host country and 200 were from outside the host country.

The results are shown in the table below.

	Predicted Winning Team		
	Brazil	**Germany**	**Other**
Host country	140	101	59
Outside the host country	72	48	80

Test at the 1% level of significance if there is an association between the location of a supporter (inside or outside the host country) and their predicted winning team.

[8 marks]

Do not write outside the box

8

Do not write outside the box

3 A random sample of size n is taken from a population which is normally distributed.

The population variance is known to be 6.25

The sample mean is used as the test statistic in a hypothesis test to investigate the population mean.

The width of a 95% confidence interval for the population mean is set at 1.4

3 (a) Find the value of n

[3 marks]

Answer _____

3 (b) A particular sample has a mean of 6.3

3 (b) (i) Find the confidence interval, giving your values to one decimal place.

[1 mark]

Answer _____

3 (b) (ii) Hence determine the conclusion of a hypothesis test at the 5% level of significance with hypotheses

$H_0: \quad \mu = 5.4$

$H_1: \quad \mu \neq 5.4$

[1 mark]

Do not write outside the box

3 (c) The true population mean is 5.7

Find the power of the test used in **part (b)(ii)**.

Give your answer to three significant figures.

[4 marks]

Answer _____

9

Do not write outside the box

4 Two independent random samples of sizes n and m

$$X_1,\ X_2,\ X_3,\ldots X_n \quad \text{and} \quad Y_1,\ Y_2,\ Y_3,\ldots Y_m$$

are taken from the same population with mean μ and variance σ^2

Unbiased estimators for σ^2 from each sample are determined as

$$S_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right) \quad \text{and} \quad S_y^2 = \frac{1}{m-1}\left(\sum_{i=1}^{m} Y_i^2 - m\overline{Y}^2\right)$$

where

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i$$

4 (a) The statistic T is given by

$$T = \frac{n}{n+m}\overline{X} + \frac{m}{n+m}\overline{Y}$$

4 (a) (i) Determine whether or not T is an unbiased estimator of μ

[3 marks]

Do not write outside the box

4 (a) (ii) Show that the variance of T is $\dfrac{\sigma^2}{n+m}$

[2 marks]

4 (a) (iii) Determine whether or not T is a consistent estimator for either sample.

[2 marks]

Question 4 continues on the next page

Do not write outside the box

4 (b) The statistic V is an unbiased estimator for σ^2 and is given by

$$V = \frac{n-1}{n+m-2}S_x^2 + kS_y^2$$

4 (b) (i) Show that $k = \frac{m-1}{n+m-2}$

[3 marks]

4 (b) (ii) State the name given to V

[1 mark]

4 (b) (iii) State a type of hypothesis test which would use V

[1 mark]

12

Do not write outside the box

Turn over for the next question

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

5 The percentage of trains that arrive at stations on time is recorded for 8 railway lines in a particular country each year.

Higher percentages for a railway line mean that more trains arrive at stations on time.

The table below shows the percentages for the 8 railway lines in 2020 and 2021.

Year	2020	2021
Line A	94.4%	97.1%
Line B	88.3%	91.1%
Line C	94.7%	94.5%
Line D	76.8%	81.0%
Line E	95.2%	93.1%
Line F	99.2%	97.5%
Line G	92.9%	94.3%
Line H	95.3%	98.4%

Test at the 10% level of significance if the percentage of trains that arrive at stations on time for the 8 railway lines has increased.

[11 marks]

Do not write outside the box

11

Turn over ►

Do not write outside the box

6 Two technology companies, A and B, manufacture silicon chips which are stored in boxes.

Both companies put the same number of silicon chips in each box.

The mean number of faulty silicon chips in a box is the same for both companies.

Company A claims that its manufacturing process is better as the variance in the number of faulty silicon chips in each box it manufactures is less than that for Company B.

The number of faulty silicon chips x in each of 12 boxes from Company A and 12 boxes from Company B is used to test this claim.

The summary statistics are shown in the table below.

	$\sum x$	$\sum x^2$
Company A	482	19 498
Company B	477	19 531

Test Company A's claim at the 1% level of significance.

[8 marks]

Do not write outside the box

8

Turn over ►

Do not write outside the box

7 The discrete random variable X has the probability distribution

x	0	1
$\mathrm{P}(X = x)$	$1 - p$	p

where $0 < p < 1$

7 (a) Show that $\mathrm{M}_X(t)$ the moment generating function of X is of the form

$$\mathrm{M}_X(t) = A + B\,\mathrm{e}^t$$

where A and B are expressions in terms of p

[2 marks]

7 (b) The discrete random variables X_i for $\{i = 1, 2, \cdots, n\}$ are independent and each have the same distribution as X above.

The sum of these variables is S

State the moment generating function for S

[1 mark]

Answer

Do not write outside the box

7 (c) (i) Use the moment generating function for S to find $\mathrm{E}(S)$ in terms of n and p

[3 marks]

Answer _____

7 (c) (ii) Use the moment generating function for S to find $\mathrm{Var}(S)$ in terms of n and p

[4 marks]

Answer _____

10

Do not write outside the box

8 A government analyst is investigating the number of successful startup companies X in each of the 400 business parks in a country.

The number of successful startup companies in each business park for a particular year are summarised in the table below.

Number of successful startup companies	0	1	2	3	4	5	6	≥ 7
Frequency	40	103	113	83	45	14	2	0

8 (a) The analyst believes that the number of successful startup companies on a business park could be modelled as a binomial distribution $X \sim \mathrm{B}(6, p)$ where p is the probability that the startup company is successful.

8 (a) (i) By calculating the mean $\overline{X}$ justify why the analyst should choose a value of $p = 0.35$

[1 mark]

8 (a) (ii) The expected frequency table below shows five values calculated using this binomial model, correct to two decimal places.

Complete the expected frequency table below using this binomial model, giving your values to two decimal places.

[2 marks]

Number of successful startup companies	0	1	2	3	4	5	6
Expected frequency	30.17	97.46	131.20		38.04	8.19	

Do not write outside the box

8 (a) (iii) Test at the 2.5% level of significance whether a binomial model is a suitable distribution for the analyst to use in this situation.

[7 marks]

Question 8 continues on the next page

Do not write outside the box

8 (b) The analyst also tries a Poisson distribution as a possible model and tests this at the 2.5% level of significance.

8 (b) (i) The expected frequency table below shows five values calculated using this Poisson model, correct to two decimal places.

Complete the expected frequency table below using this Poisson model, giving your values to two decimal places.

[2 marks]

Number of successful startup companies	0	1	2	3	4	5	≥ 6
Expected frequency		102.86	108.01	75.60	39.69		8.18

8 (b) (ii) The test statistic for the data is 8.41

Find the critical value and state the conclusion of the test.

[3 marks]

Critical Value

Conclusion

8 (c) Using the results of your tests in **part (a)(iii)** for the binomial model and in **part (b)(ii)** for the Poisson model state with a reason which is the better model for successful startup companies in business parks.

[1 mark]

16

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.**
	Copyright information For confidentiality purposes, all acknowledgements of third-party copyright material are published in a separate booklet. This booklet is published after each live examination series and is available for free download from www.oxfordaqaexams.org.uk. Permission to reproduce all copyright material has been applied for. In some cases, efforts to contact copyright-holders may have been unsuccessful and Oxford International AQA Examinations will be happy to rectify any omissions of acknowledgements. If you have any queries please contact the Copyright Team.