

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.

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	
TOTAL	



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

$$H_0: \mu_Y = \mu_X$$

$$H_1: \mu_Y > \mu_X$$

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]



- 4 Two independent random samples of sizes n and m

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

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

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

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

where

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

- 4 (a) The statistic T is given by

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

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

[3 marks]



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

[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

Turn over ►



Turn over for the next question

*Do not write
outside the
box*

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**

Turn over ►



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]



7 The discrete random variable X has the probability distribution

x	0	1
$P(X = x)$	$1 - p$	p

where $0 < p < 1$

7 (a) Show that $M_X(t)$ the moment generating function of X is of the form

$$M_X(t) = A + B e^t$$

where A and B are expressions in terms of p

[2 marks]

7 (b) The discrete random variables X_i for $\{i = 1, 2, \dots, n\}$ 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 _____



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

[3 marks]

Answer _____

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

[4 marks]

Answer _____

10

Turn over ►



- 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 $X \sim B(6, p)$ where p is the probability that the startup company is successful.

- 8 (a) (i)** By calculating the mean \bar{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	



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]

END OF QUESTIONS



There are no questions printed on this page

*Do not write
outside the
box*

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**



