Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number

Candidate Number

Pearson Edexcel Level 3 GCE

Paper reference **8FM0/23**

Further Mathematics

Advanced Subsidiary
Further Mathematics options
23: Further Statistics 1
(Part of options B, E, F and G)

You must have:
Mathematical Formulae and Statistical Tables (Green), calculator

Total Marks

Candidates may use any calculator allowed by Pearson regulations. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. If a calculator is used instead of the tables, the value should be given to an equivalent degree of accuracy.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- The total mark for this part of the examination is 40. There are 4 questions.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ▶

1. Stuart is investigating a treatment for a disease that affects fruit trees. He has 400 fruit trees and applies the treatment to a random sample of these trees. The remainder of the trees have no treatment. He records the number of years, y, that each fruit tree remains free from this disease.

The results are summarised in the table below.

		Treatment		
		Applied	Not applied	
Number of years free from this disease	$y < 1$	15	25	40
	$1 \leqslant y < 2$	35	61	96
	$2 \leqslant y$	124	140	264
		174	226	400

The data are to be used to determine whether or not there is an association between the application of the treatment and the number of years that a fruit tree remains free from this disease.

(a) Calculate the expected frequencies for

(i) Applied and $y < 1$

(ii) Not applied and $1 \leqslant y < 2$

(2)

The value of $\sum \frac{(O-E)^2}{E}$ for the **other four** classes is 2.642 to 3 decimal places.

(b) Test, at the 5% level of significance, whether or not there is an association between the application of the treatment and the number of years a fruit tree remains free from this disease.

You should state your hypotheses, test statistic, critical value and conclusion clearly.

(5)

y>1 ↓ applied ↓

(a) (i) $\frac{40 \times 174}{400} = 17.4$ ← using row + column totals. (annotated on table in Q1)

(ii) $\frac{96 \times 226}{400} = 54.24$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 1 continued

(b) H_0: There is no association between the application of the treatment and the number of years that a fruit tree remains free from disease.

H_1: There is an association between the application of the treatment and the number of years that a fruit tree remains free from disease.

① for correct H_0 AND H_1

$$\sum \frac{(O-E)^2}{E} = \frac{(15-17.4)^2}{17.4} + \frac{(61-54.24)^2}{54.24} + 2.642 \quad ①$$

$$= 3.82 \ (3 \text{ s.f}) \quad ①$$

Degrees of freedom $= 2$ ← (rows − 1) × (columns − 1) = 2 × 1

$\chi^2_2(5\%) = 5.991$ ← critical value read from statistical tables (in formulae book)

↑ significance level = 5% = 0.05

$3.82 < 5.991$ ① ∴ There is no evidence of an association between the application of the treatment and the number of years that a fruit tree remains free from disease ① (Accept H_0).

Question 1 continued

Question 1 continued

(Total for Question 1 is 7 marks)

2. Xena catches fish at random, at a constant rate of 0.6 per hour.

(a) Find the probability that Xena catches exactly 4 fish in a 5-hour period.

(2)

The probability of Xena catching no fish in a period of t hours is less than 0.16

(b) Find the minimum value of t, giving your answer to one decimal place.

(3)

Independently of Xena, Zion catches fish at random with a mean rate of 0.8 per hour.

Xena and Zion try using new bait to catch fish. The number of fish caught in total by Xena and Zion after using the new bait, in a randomly selected 4-hour period, is 12

(c) Use a suitable test to determine, at the 5% level of significance, whether or not there is evidence that the rate at which fish are caught has increased after using the new bait. State your hypotheses clearly and the *p*-value used in your test.

(5)

(a) $X \sim Po(3)$ ① ← use a Poisson distribution for rates (events per time period)

$P(X = 4) = 0.1680$ ① ← put in calculator or use $\frac{e^{-\lambda} \times \lambda^{x}}{x!}$

(b) $X \sim Po(0.6t)$

$P(X = 0) > \frac{e^{-\lambda} \times \lambda^{x}}{x!}$

$0.16 > \frac{e^{-0.6t} \times 0.6t^{0}}{1}$ ①

$0.16 > e^{-0.6t}$

$-0.6 \times t < \ln(0.16)$ ①

$t > 3.05\ldots$

$t = 3.1$ ①

DO NOT WRITE IN THIS AREA

Question 2 continued

(c) $H_0 : \lambda = 1.4$ $H_1 : \lambda > 1.4$ (1) ← 0.6 + 0.8 per hour

$J \sim Po(5.6)$ (1) ← 1.4 × 4 hours

$P(J \geq 12) = 1 - P(J \leq 11)$ (1)

$= 1 - 0.9875$

$= 0.01248$ (1) ← (p-value)

$0.01248 < 0.05$ ∴ Reject H_0. ← 5% significance = 0.05

There is evidence at the 5% level of significance that the rate of fish caught may have increased. (1)

Question 2 continued

Question 2 continued

(Total for Question 2 is 10 marks)

3. In a game, a coin is spun 5 times and the number of heads obtained is recorded. Tao suggests playing the game 20 times and carrying out a chi-squared test to investigate whether the coin might be biased.

(a) Explain why playing the game only 20 times may cause problems when carrying out the test.

(1)

Chris decides to play the game 500 times. The results are as follows

Number of heads	0	1	2	3	4	5
Observed frequency	2	27	93	181	146	51

Chris decides to test whether or not the data can be modelled by a binomial distribution, with the probability of a head on each spin being 0.6

She calculates the expected frequencies, to 2 decimal places, as follows

Number of heads	0	1	2	3	4	5
Expected frequency	5.12	38.40	115.20	172.80	129.60	38.88

(b) State the number of degrees of freedom in Chris' test, giving a reason for your answer.

(1)

(c) Carry out the test at the 5% level of significance.
You should state your hypotheses, test statistic, critical value and conclusion clearly.

(5)

(d) Showing your working, find an alternative model which would better fit Chris' data.

(2)

(a) Sample size is too small ①

(b) 5 degrees of freedom because the parameter is not being estimated from the data, and the totals agree. ①

DO NOT WRITE IN THIS AREA

Question 3 continued

(c) H_0 : B(5, 0.6) is a suitable model

① for correct H_0 AND H_1

H_1 : B(5, 0.6) is not a suitable model.

$$\sum \frac{(O-E)^2}{E} = \frac{(2-5.12)^2}{5.12} + \frac{(1-38.40)^2}{38.40} + \frac{(2-115.20)^2}{115.20} + \frac{(3-172.8)^2}{172.8} + \frac{(4-129.6)^2}{129.6} + \frac{(5-38.88)^2}{38.88}$$ ①

$= 15.8063$. ①

$\chi^2_5(5\%) = 11.070$ ① ← from statistical tables in formulae book

$15.8 > 11.07$ ∴ Reject H_0

There is evidence at the 5% significance level to suggest that B(5, 0.6) is not a suitable model. ①

(d) $\frac{0 \cdot 2\times 0 + 1\times 27 + 2\times 93 + 3\times 181 + 4\times 146 + 5\times 51}{500} = 3.19$ ①

↑ using actual observed outcomes

$p = \frac{3.19}{5} = 0.638$ ← $\frac{\text{frequency}}{\text{no. of trials}}$

∴ A better model is B(5, 0.638) ①

Question 3 continued

Question 3 continued

(Total for Question 3 is 9 marks)

4. The discrete random variable X has the following probability distribution

x	0	2	3	6
$P(X = x)$	p	0.25	q	0.4

(a) Find in terms of q

(i) $E(X)$

(ii) $E(X^2)$

(2)

Given that $\text{Var}(X) = 3.66$

(b) show that $q = 0.3$

(3)

In a game, the score is given by the discrete random variable X

Given that games are independent,

(c) calculate the probability that after the 4th game has been played, the total score is exactly 20

(3)

A round consists of 4 games plus 2 bonus games. The bonus games are only played if after the 4th game has been played the total score is exactly 20

A prize of £10 is awarded if 6 games are played in a round **and** the total score for the round is at least 27

Bobby plays 3 rounds.

(d) Find the probability that Bobby wins at least £10

(6)

(a)(i) $E(X) = \sum_{i=1}^{n} x_i \times P(X = x)$

$= 0 \times p + 2 \times 0.25 + 3 \times q + 6 \times 0.4$

$= 2.9 + 3q$ ①

(ii) $E(X^2) = \sum_{i=1}^{n} x_i^2 \times P(X = x)$

$= 0^2 \times p + 2^2 \times 0.25 + 3^2 \times q + 6^2 \times 0.4$

$= 15.4 + 9q$ ①

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 4 continued

(b) $Var(X) = E(X^2) - E(X)^2$

$3.66 = (15.4 + 9q) - (2.9 + 3q)^2$ ①

$3.66 = 15.4 + 9q - 8.41 - 9q^2 - 17.4q$

$9q^2 + 8.4q - 3.33 = 0$

$q = 0.3$ and $q = -\frac{37}{30}$ ①

$\therefore q = 0.3$ because $q \geq 0$ ①

(c) $P(x_1 + x_2 + x_3 + x_4 = 20) = P(6,6,6,2$ or $6,6,2,6$ or $6,2,6,6$ or $2,6,6,6)$ ①

$= 4 \times 0.4^3 \times 0.25$ ①

↑ score = 6 for 3 games; ↑ score = 2 for 1 game

$= 0.064$ ①

(d) $P(x_5 + x_6 \geq 7) = P(6,6$ or $6,3$ or $6,2)$ ①

$= 0.4^2 + 2 \times (0.4 \times 0.3) + 2 \times (0.4 \times 0.25)$

$= 0.6$ ①

$P(\text{score} \geq 27) = 0.064 \times 0.6$ ← P(bonus played AND score ≥ 7)

$= 0.0384$ ①

Question 4 continued

$Y \sim B(3, 0.0384)$ (1) ← binomial because he either wins OR he doesn't ∴ binary outcome.

$P(Y \geq 1) = 1 - P(Y = 0)$ (1)

$= 1 - 0.8892$

$= 0.1108$ (1)

(Total for Question 4 is 14 marks)

TOTAL FOR FURTHER STATISTICS 1 IS 40 MARKS