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 **8FM0/24**

Further Mathematics

Advanced Subsidiary
Further Mathematics options
24: Further Statistics 2
(Part of option G only)

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 statistical tables should be quoted in full. If a calculator is used instead of 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 5 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. Abena and Meghan are both given the same list of 10 films.

Each of them ranks the 10 films from most favourite to least favourite.

For the differences, d, between their ranks for these 10 films, $\sum d^2 = 84$

(a) Calculate Spearman's rank correlation coefficient between Abena's ranks and Meghan's ranks.

(1)

A test is carried out at the 5% level of significance to see if there is agreement between their ranks for the films.

The hypotheses for the test are

$$\mathrm{H}_0 : \rho_S = 0 \qquad \mathrm{H}_1 : \rho_S > 0$$

(b) (i) Find the critical region for the test.

(ii) State the conclusion of the test.

(2)

An 11th film is added to the list. Abena and Meghan both agree that this film is their least favourite.

A new test is carried out at the 5% level of significance using the same hypotheses.

(c) Determine the conclusion of this test. You should state the test statistic and the critical value used.

(4)

(a) $r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$ $\Sigma d^2 = 84$

$n = 10$ (pairs of observations)

$r_s = 1 - \frac{6 \times 84}{10(10^2-1)} = 0.491$ (3.s.f) 1

(b)(i) 5% = 0.05 significance level, sample size = 10

Use table in formula book (for r_s) to find critical value

Critical val : 0.5636 1 $\Rightarrow$ Critical region : $0.5636 < r_s < 1$

(ii) 0.491 is not in the critical region, so there is insufficient evidence of agreement between their film rankings. 1

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

Question 1 continued

(c) New $r_s = 1 - \frac{6 \times 84}{11(11^2 - 1)}$ (1) $= 0.618$ (1)

At $5\% = 0.05$ sig. level, sample size $= 11$:

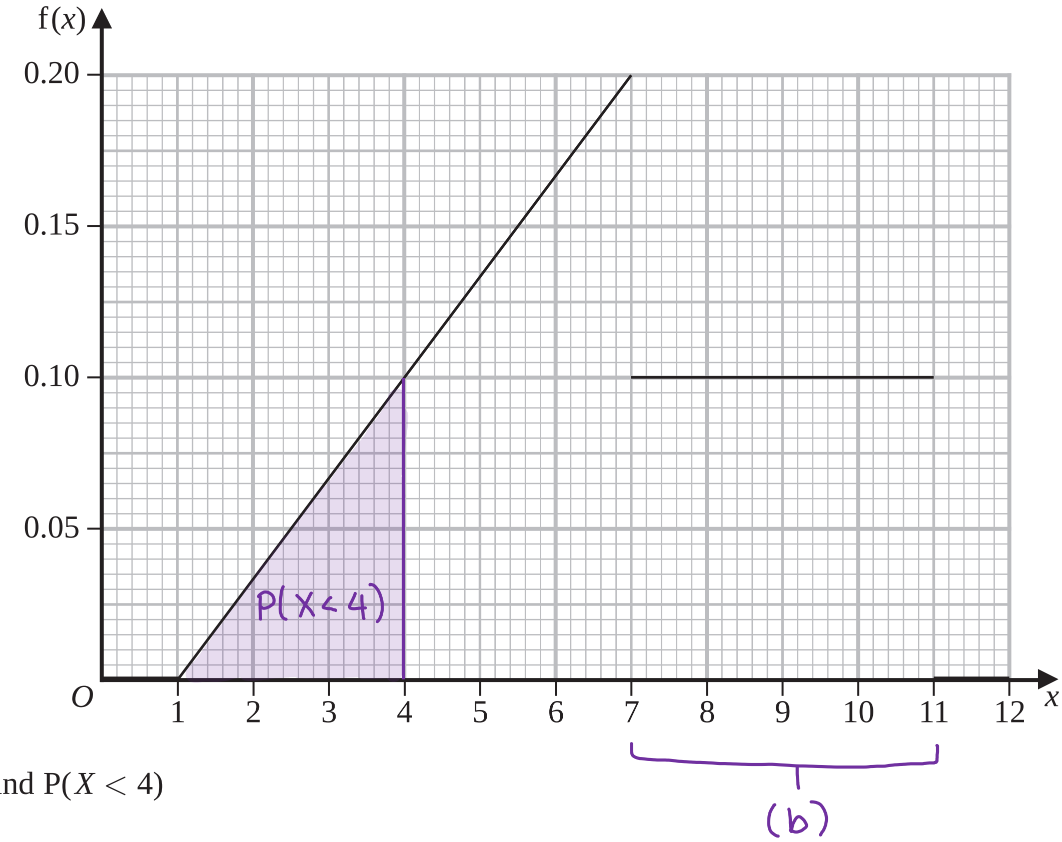
crit. val $= 0.5364$ (1) ← using same stat. table.

crit. region: $0.5364 < r_s < 1$

Test statistic (0.618) falls in critical region, so there is now sufficient evidence of agreement between their film rankings. (1)

(Total for Question 1 is 7 marks)

2. The graph shows the probability density function $f(x)$ of the continuous random variable X



(a) Find $P(X < 4)$

(2)

(b) Specify the cumulative distribution function of X for $7 \leqslant x \leqslant 11$

(3)

(a) P = area under function between given limits.

$P(X < 4) = \frac{(4-1)\times 0.1}{2}$ ① $= 0.15$ ①

(b) $F(x) = P(X \leq x)$ ← equivalent to integrating between the lower limit and x

$F(x) = \int 0.1\,dx - P(X \leq 7)$

$F(x) = 0.1x + c - \frac{(7-1)\times 0.2}{2}$ ①

$F(x) = 0.1x + c - 0.6$

Question 2 continued

$F(7) = 0 \Rightarrow 0.1(7) + c - 0.6 = 0$

$0.7 + c - 0.6 = 0$

$\therefore c = -0.1$ ①

$F(x) = 0.1x - 0.1$ for $7 \leq x \leq 11$ ①

(Total for Question 2 is 5 marks)

3. Gabriela is investigating a particular type of fish, called bream. She wants to create a model to predict the weight, w grams, of bream based on their length, x cm.

For a sample of 27 bream, some summary statistics are given below.

$$\bar{x} = 31.07 \qquad \bar{w} = 628.59 \qquad \sum w^2 = 11\,386\,134$$

$$S_{xw} = 13\,082.3 \qquad S_{xx} = 260.8$$

(a) Find the value of the product moment correlation coefficient between x and w

(3)

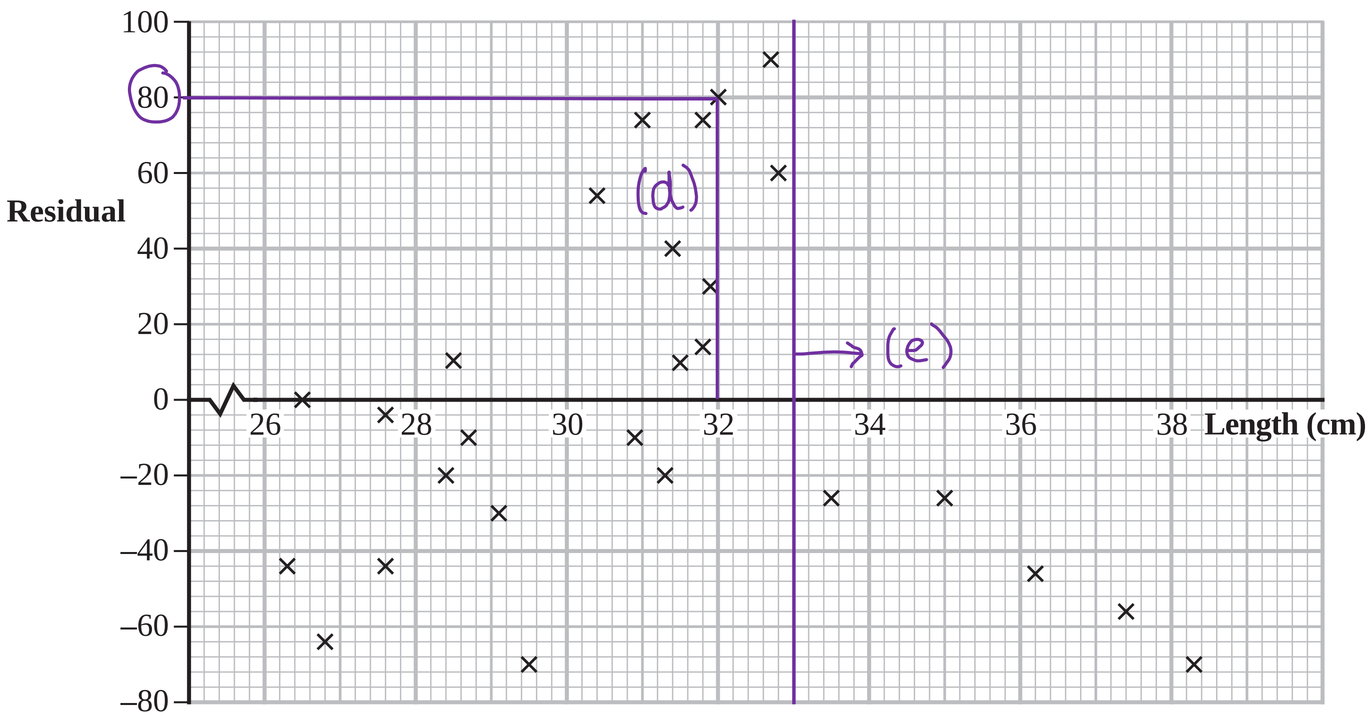
(b) Explain whether the answer to part (a) is consistent with a linear model for these data.

(1)

(c) Find the equation of the regression line of w on x in the form $w = a + bx$

(3)

A residual plot for these data is shown below.



One of the bream in the sample has a length of 32 cm.

(d) Find its weight.

(2)

(e) With reference to the residual plot, comment on the model for bream with lengths above 33 cm.

(1)

Question 3 continued

(a) $\text{PMCC} = \dfrac{S_{xw}}{\sqrt{S_{xx} \times S_{ww}}}$ need to find S_{ww}

$S_{ww} = \Sigma w^2 - \dfrac{(\Sigma w)^2}{n}$

$\bar{w} \times \text{sample size} = \Sigma w$

$S_{ww} = 11\,386\,134 - 27(628.59)^2 = 717\,748.5213$ 1

$\text{PMCC} = \dfrac{13\,082\ 3}{\sqrt{260\ 8 \times 717\,748.5213}}$ 1 $= 0.956$ 1 (3.s.f)

(b) Since PMCC is close to 1, the data is consistent with a linear model. 1

(c) $b = \dfrac{S_{xw}}{S_{xx}} = \dfrac{13\,082.3}{260\ 8} = 50.2$ 1

$a = \bar{w} - b\bar{x} = 628.59 - 50.162(31\ 07) = -930$ 1

$w = a + bx \Rightarrow w = -930 + 50\ 2x$ 1

(d) From graph, residual at 32cm = +80

$w = -930 + 50\ 2(32) + 80$ 1 $= 756\ 4$ 1

(e) Negative residuals for all $x > 33$ suggests the model systematically overestimates weights for the longer bream. 1

Question 3 continued

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

Question 3 continued

(Total for Question 3 is 10 marks)

4. A random variable X has probability density function given by

$$f(x) = \begin{cases} 0.8 - 6.4x^{-3} & 2 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

The median of X is m

(a) Show that $m^3 - 3.625m^2 + 4 = 0$ **(3)**

(b) (i) Find $f'(x)$

(ii) Explain why the mode of X is 4 **(2)**

Given that $E(X^2) = 10.5$ to 3 significant figures,

(c) find $\text{Var}(X)$, showing your working clearly. **(4)**

(a) Median is such that $F(m) = 0.5$

$F(m) = P(X \leq m)$

$F(m) = \int_2^m (0.8 - 6.4x^{-3})\,dx = 0.5$ 1

$0.5 = \left[0.8x + 3.2x^{-2}\right]_2^m$ 1

$0.5 = \left[0.8(m) + 3.2(m)^{-2}\right] - \left[0.8(2) + 3.2(2)^{-2}\right]$

$0.5 = 0.8m + \frac{3.2}{m^2} - 2.4$

$0 = 0.8m + \frac{3.2}{m^2} - 2.9$ $\times m^2$

$0 = 0.8m^3 - 2.9m^2 + 3.2$ $\div 0.8$

$0 = m^3 - 3.625m^2 + 4$ 1

(b)(i) $f(x) = 0.8 - 6.4x^{-3}$

$f'(x) = -3 \times -6.4x^{-3-1} = 19.2x^{-4}$ 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 DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Question 4 continued

(b)(ii) Mode of continuous random variable is value of x that yields the maximum of the p.d.f.

Since $f'(x) > 0$, $f(x)$ is increasing so the p.d.f has its maximum value at the upper end of the interval. Therefore the mode is 4

upper end of $2 \leq x \leq 4$

(c) $\text{Var}(X) = E(X^2) - [E(X)^2]$

$E(X) = \sum x \times P(X \leq x) = \int_a^b x f(x)\,dx$

$\text{Var}(X) = 10.5 - \left[\int_2^4 x(0.8 - 6.4x^{-3})\,dx\right]^2$ 1

$\text{Var}(X) = 10.5 - \left[\left[0.4x^2 + 6.4x^{-1}\right]_2^4\right]^2$

$\text{Var}(X) = 10.5 - \left[0.4(4^2) + \frac{6.4}{4} - 0.4(2^2) - \frac{6.4}{2}\right]^2$ 1

$\text{Var}(X) = 10.5 - 3.2^2$ 1

$\text{Var}(X) = 0.26$ 1

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 4 continued

(Total for Question 4 is 9 marks)

5. The random variable X has the continuous uniform distribution over the interval [0.5, 2.5]

Talia selects a number, T, at random from the distribution of X

(a) Find $P(T < 1)$

(1)

Malik takes Talia's number, T, and calculates his number, M, where $M = \frac{1}{T^2}$

(b) Find the probability that both T and M are less than 2.25

(3)

Raja and Greta play a game many times.
Each time they play they use a number, R, randomly selected from the distribution of X

Raja's score is R

Greta's score is G, where $G = \frac{2}{R^2}$

(c) Determine, giving a reason, who you would expect to have the higher **total score**.

(5)

(a) $X \sim U[0.5, 2.5]$ ← uniform, so any value in the range is equally likely.

$P(T<1) = \frac{1-0.5}{2.5-0.5}$ ← length where $T<1$ ← total length of interval

$P(T<1) = \frac{1}{4}$ (1)

(b) $P(\{T < 2.25\} \cap \{\frac{1}{T^2} < 2.25\})$

$= P(\{T < 2.25\} \cap \{T^2 > \frac{4}{9}\})$ (1)

$P(\frac{2}{3} < T < 2.25) = \frac{2.25 - \frac{2}{3}}{2.5 - 0.5}$ (1) $= \frac{19}{24}$ (1)

$\sqrt{\frac{4}{9}} = \frac{2}{3}$

DO NOT WRITE IN THIS AREA

Question 5 continued

(c) $E(R) = \frac{0.5 + 2.5}{2} = 1.5$ ① ← Raja's expected total score

p.d.f $f(x) = \begin{cases} \frac{1}{2.5-0.5} & \text{for } 0.5 \le x \le 2.5 \\ 0 & \text{otherwise} \end{cases}$

$\therefore E\left(\frac{2}{R^2}\right) = \int_{0.5}^{2.5} \frac{1}{2.5-0.5} \times \frac{2}{r^2}\, dr$ ①

$= \int_{0.5}^{2.5} \frac{1}{r^2}\, dr$

$= \left[-\frac{1}{r}\right]_{0.5}^{2.5}$ ①

$= -\frac{1}{2.5} + \frac{1}{0.5}$

$= 1.6$ ① ← Greta's expected total score

$\therefore$ Greta is the expected winner since she has the higher expected value $(1.6 > 1.5)$. ①

P 7 1 9 6 8 A 0 1 5 1 6

Question 5 continued

(Total for Question 5 is 9 marks)

TOTAL FOR FURTHER STATISTICS 2 IS 40 MARKS