

Please check the examination details below before entering your candidate information

Candidate surname

Other names

**Pearson Edexcel
Level 3 GCE**

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

Thursday 08 October 2020

Afternoon

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 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 ►

P62659A

©2020 Pearson Education Ltd.

1/1/1/1/



Pearson

1. An estate agent in *Tornep* believes that houses further from the railway station are more expensive than those that are closer. She took a random sample of 22 three-bedroom houses in *Tornep* and calculated the product moment correlation coefficient between the house price and the distance from the station to be 0.3892

Stating your hypotheses clearly, use a 5% level of significance to test the estate agent's belief. State the critical region used in your test.

1. state hypotheses (3)

$H_0: \rho = 0$ $H_1: \rho > 0$

larger distance \leftrightarrow
more expensive
 \hookrightarrow +ve correlation

2. find critical value

from tables, 5% one-tail critical value for r is

0.3598

3. compare values & form conclusion

$0.3892 > 0.3598 \rightarrow$ significant result so there is

sufficient evidence to support the agent's belief.



Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 1 is 3 marks)



P 6 2 6 5 9 A 0 3 1 6

2. Mary, Jahil and Dawn are judging the cakes in a village show. They have 5 features to consider and each feature is awarded up to 5 points. The total score the judges gave each cake are given in the table below.

Cake	A	B	C	D	E	F	G	H	I
Mary	19	17	23	10	21	15	12	8	14
Jahil	22	18	21	10	24	20	16	12	15
Dawn	9	11	6	18	9	15	13	20	13

- (a) Calculate Spearman's rank correlation coefficient between Mary's scores and Jahil's scores. (4)

- (b) Calculate Spearman's rank correlation coefficient between Jahil's scores and Dawn's scores. (3)

The judges discussed their interpretation of the points system and agreed that the first prize should go to cake C.

- (c) Explain how different interpretations of the points system could give rise to the results in part (a) and part (b). (2)

a) rank M & J's scores highest to lowest:

Cake	A	B	C	D	E	F	G	H	I
M	3	4	1	8	2	5	7	9	6
J	2	5	3	9	1	4	6	8	7

square of differences between M & J:

$$\sum d^2 = 1 + 1 + 4 + 1 + 1 + 1 + 1 + 1 + 1 = 12$$

$$\text{no tied ranks} \Rightarrow r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$n=9: r_s = 1 - \frac{6 \times 12}{9 \times (81 - 1)} = \underline{0.9}$$



Question 2 continued

b) rank J & D's scores highest to lowest:

Cake	A	B	C	D	E	F	G	H	I
J	2	5	3	9	1	4	6	8	7
D	7.5	6	9	2	7.5	3	4.5	1	4.5

equal data values
given rank = mean
of tied ranks

multiple ties \Rightarrow use PMCC on ranks

$$S_{JJ} = \sum r_J^2 - \frac{(\sum r_J)^2}{9} = 285 - \frac{45^2}{9} = 60$$

$$S_{DD} = \sum r_D^2 - \frac{(\sum r_D)^2}{9} = 284 - \frac{45^2}{9} = 59$$

$$S_{JD} = 176 - \frac{45^2}{9} = -49$$

$$\Rightarrow r_s = \frac{-49}{\sqrt{60 \times 59}} = -0.8236 \quad (4sf)$$

c) possibly, M&J awarded higher points for good features

whereas D gave more points for poor features.

there is a strong correlation in both cases, with M&J +ve

& J&D -ve, so the judges agree which cake is best

overall.



Question 2 continued

Lined writing area for the answer to Question 2.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 2 is 9 marks)



P 6 2 6 5 9 A 0 7 1 6

3. The continuous random variable X has cumulative distribution function

$$F(x) = \begin{cases} 0 & x < 4 \\ px - k\sqrt{x} & 4 \leq x \leq 9 \\ 1 & x > 9 \end{cases}$$

where p and k are constants.

- (a) Find the value of p and the value of k .

(4)

Given that $E(X) = \frac{119}{18}$

- (b) show that $\text{Var}(X) = 2.05$ to 3 significant figures.

(6)

- (c) Write down the mode of X .

(1)

- (d) Find the exact value of the constant a such that $P(X \leq a) = \frac{7}{27}$

(3)

$$a) F(4) = 0 \Rightarrow 4p - 2k = 0$$

$$F(9) = 1 \Rightarrow 9p - 3k = 1$$

} distribution is smooth \Rightarrow no jump
@ $x=4, x=9$

$$k = 2p$$

$$\text{sub in: } 9p - 6p = 1$$

$$3p = 1$$

$$p = \frac{1}{3}$$

$$\therefore k = \frac{2}{3}$$

differentiate
cumulative d.f. \rightarrow p.d.f.

$$b) \text{Var}(X) = E(X^2) - (E(X))^2, \text{ so we need } E(X)$$

$$\text{need } f(x): f(x) = F'(x) = \frac{1}{3} - \frac{2}{3} \times \frac{1}{2} x^{-\frac{1}{2}} = \frac{1}{3} (1 - x^{-\frac{1}{2}})$$

$$\text{so } E(X^2) = \frac{1}{3} \int_4^9 x^2 (1 - x^{-\frac{1}{2}}) dx = \frac{1}{3} \int_4^9 (x^2 - x^{\frac{3}{2}}) dx$$

range where
 $f(x) \neq 0$



Question 3 continued

$$\begin{aligned} &= \frac{1}{3} \left[\frac{x^3}{3} - \frac{2x^{\frac{5}{2}}}{5} \right]_4^9 \\ &= \frac{1}{3} \left[\left(\frac{9^3}{3} - \frac{2 \times 3^5}{5} \right) - \left(\frac{64}{3} - \frac{2 \times 2^5}{5} \right) \right] \\ &= \frac{2059}{45} \end{aligned}$$

$$\begin{aligned} \text{so } \text{Var}(X) &= \frac{2059}{45} - \left(\frac{119}{18} \right)^2 \\ &= 2.048765... \\ &= 2.05 \quad (3 \text{ s.f.}) \end{aligned}$$

c) $f(x) = \frac{1}{3} \left(1 - \frac{1}{\sqrt{x}} \right) \therefore$ max. when x is greatest ($\frac{1}{\sqrt{x}}$ smallest)

x -range for $f(x) \neq 0$ is $4 \leq x \leq 9 \therefore$ mode = 9

$$\text{d) } P(X \leq a) = F(a) = \frac{7}{27}$$

$$\Rightarrow \frac{1}{3} \cdot a - \frac{2\sqrt{a}}{3} = \frac{7}{27}$$

$$a - 2\sqrt{a} - \frac{7}{9} = 0 \quad (\times 3)$$

$$\text{use } y = \sqrt{a} : y^2 - 2y - \frac{7}{9} = 0$$



Question 3 continued

$$\Rightarrow y = \frac{2 \pm \sqrt{2^2 - 4(-\frac{7}{9})}}{2} = \frac{7}{3} \text{ or } -\frac{1}{3}$$

root of a -ve number not valid $\therefore y = \frac{7}{3}, a = \frac{49}{9}$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 3 is 14 marks)



P 6 2 6 5 9 A 0 1 1 1 6

4. Some students are investigating the strength of wire by suspending a weight at the end of the wire. They measure the diameter of the wire, d mm, and the weight, w grams, when the wire fails. Their results are given in the following table.

These 14 points are plotted on page 13														Not yet plotted				
d	0.5	0.6	0.7	0.8	0.9	1.1	1.3	1.6	2	2.4	2.8	3.3	3.5	3.9	4.5	4.6	4.8	5.4
w	1.2	1.7	2.3	3.0	3.8	5.6	7.7	11.6	18	25.9	34.9	47.4	52.7	63.9	81	83.6	89.9	109.4

The first 14 points are plotted on the axes on page 13.

- (a) On the axes on page 13, complete the scatter diagram for these data. (1)
- (b) Use your calculator to write down the equation of the regression line of w on d . (2)
- (c) With reference to the scatter diagram, comment on the appropriateness of using this linear regression model to make predictions for w for different values of d between 0.5 and 5.4 (1)

The product moment correlation coefficient for these data is $r = 0.987$ (to 3 significant figures).

- (d) Calculate the residual sum of squares (RSS) for this model. (2)

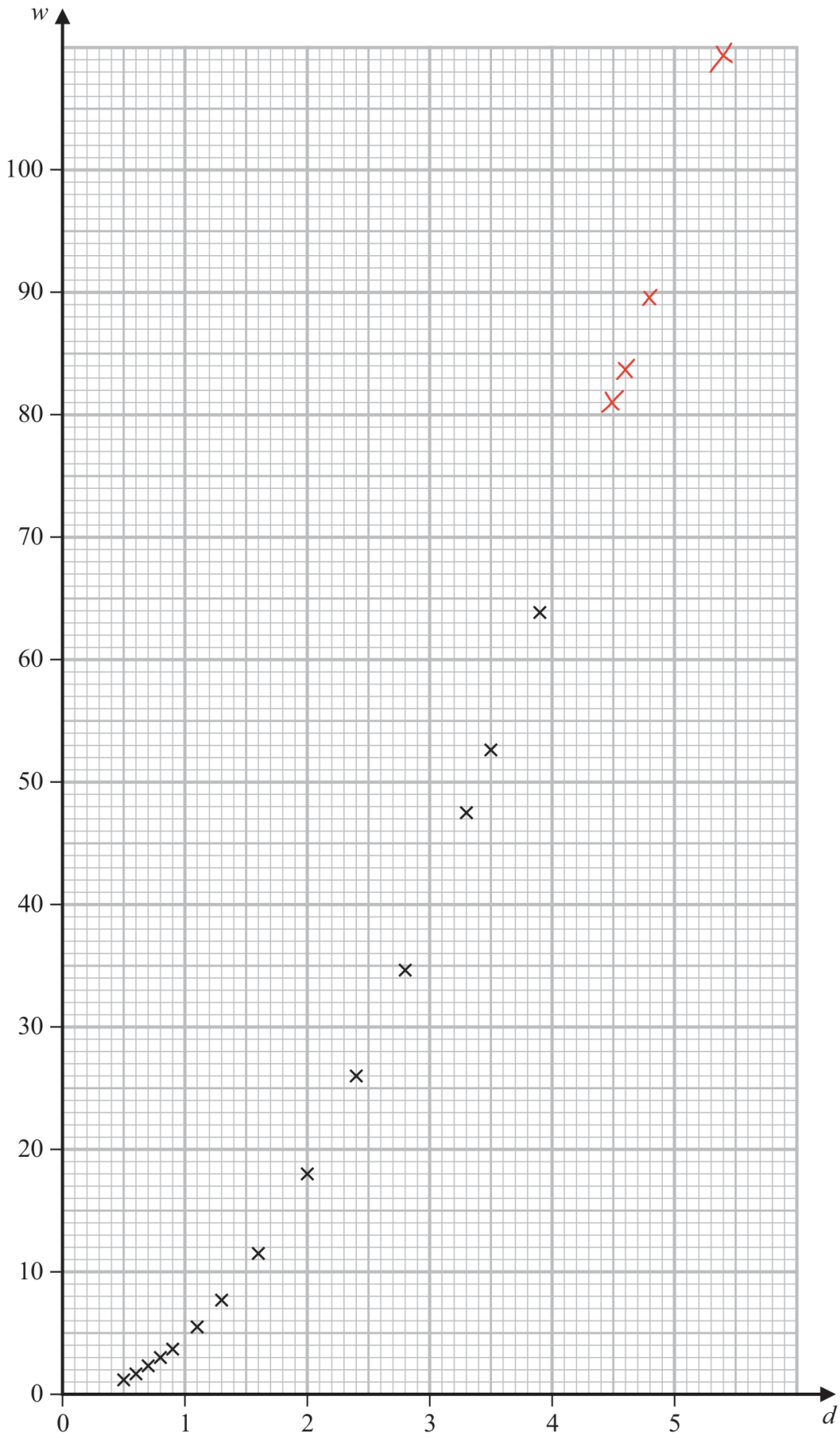
Robert, one of the students, suggests that the model could be improved and intends to find the equation of the line of regression of w on u , where $u = d^2$. He finds the following statistics

$$S_{wu} = 5721.625 \quad S_{uu} = 1482.619 \quad \sum u = 157.57$$

- (e) By considering the physical nature of the problem, give a reason to support Robert's suggestion. (1)
- (f) Find the equation of the regression line of w on u . (3)
- (g) Find the residual sum of squares (RSS) for Robert's model. (2)
- (h) State, giving a reason based on these calculations, which of these models better describes these data. (1)
- (i) Hence estimate the weight at which a piece of wire with diameter 3 mm will fail. (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

b) model of form $w = a + bd$ ('line of w on d ')

$$\underline{w = 21.5d - 17.7}$$

c) this wouldn't be appropriate \because the plot suggests a more curved trend, so a linear model would overestimate in the middle & underestimate elsewhere.

$$\begin{aligned} d) S_{ww} &= \sum w^2 - \frac{(\sum w)^2}{18} \\ &= 45178.68 - \frac{643.6^2}{18} \\ &= 22166.404... \end{aligned}$$

$$\begin{aligned} RSS &= S_{ww} (1 - r^2) = 22166.404... \times (1 - 0.987^2) \\ &= 570 \text{ g}^2 \end{aligned}$$

e) thicker wire should be stronger so strength is proportional to area, which is $\propto d^2$

f) 'w on u': $w = cu + f$

$$c = \frac{S_{wu}}{S_{uu}} = \frac{5721.625}{1482.619} = 3.85913$$



Question 4 continued

$$f = \bar{w} - c\bar{u}$$

$$= \frac{\sum w}{18} - c \frac{\sum u}{18}$$

$$= \frac{643.6}{18} - 3.8591 \times \frac{157.57}{18}$$

$$= 1.973\dots$$

$$\text{so } w = 1.97 + 3.86u$$

$$g) \text{ RSS} = S_{ww} - \frac{(S_{wu})^2}{S_{uu}}$$

$$= 22166.404 - \frac{5721.625^2}{1482.619}$$

↑
from (d)

$$= 85.88\dots$$

$$= 85.9 \text{ g}^2 \text{ (3 s.f.)}$$

h) Robert's model has a lower RSS so is better

this means each
point is closer to
the model compared
with the initial one

$$i) u = d^2, w = 3.86u + 1.97$$

$$\hookrightarrow w = 3.86 \times 3^2 + 1.97 = 36.71 \text{ g} = 36.7 \text{ g (3sf)}$$



Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 4 is 14 marks)

**TOTAL FOR FURTHER STATISTICS 2 IS 40 MARKS
END**

