Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Thursday 16 May 2019

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.
- 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 ▶

P61861A

©2019 Pearson Education Ltd.

1/1/1/1/

P 6 1 8 6 1 A 0 1 1 6

Pearson

stemup.co.uk

1. Bara is investigating whether or not the two judges of a skating competition are in agreement. The two judges gave a score to each of the 8 skaters in the competition as shown in the table below.

	Skater							
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Judge 1	71	70	72	62	63	61	57	53
Judge 2	73	71	67	64	62	56	52	53

Bara decided to calculate Spearman's rank correlation coefficient for these data.

(a) Calculate Spearman's rank correlation coefficient between the ranks of the two judges.

(4)

(b) Test, at the 1% level of significance, whether or not the two judges are in agreement.

(4)

Judge 1 accidentally swapped the scores for skaters *D* and *E*. The score for skater *D* should be 63 and the score for skater *E* should be 62

(c) Without carrying out any further calculations, explain how Spearman's rank correlation coefficient will change. Give a reason for your answer.

(2)

(a)

Skater	A	B	C	D	E	F	G	H
Judge 1	2	3	1	5	4	6	7	8
Judge 2	1	2	3	4	5	6	8	7
d	1	1	−2	1	−1	0	−1	1
d^2	1	1	4	1	1	0	1	1

$\Sigma d^2 = 1 + 1 + 4 + 1 + 1 + 0 + 1 + 1 = 10$

$r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6 \times 10}{8(64-1)} = 1 - \frac{60}{504} = 0.880952...$

$r_s = 0.881$ (3 s.f.)

(b) H_0: $\rho = 0$ H_1: $\rho > 0$

$n = 8$, 1% one-tailed critical value = 0.8333

$0.881 > 0.8333$, so r_s lies in the critical region: reject H_0

There is evidence that the two judges are in agreement.

(c) With the corrected scores Judge 1 ranks D as 4th and E as 5th, which is the same as Judge 2 (previously they were different), so $d = 0$ for D and E.

Σd^2 decreases (by 2, from 10 to 8), so Spearman's rank correlation coefficient will increase (become closer to 1).

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

stemup.co.uk

Question 1 continued

(Total for Question 1 is 10 marks)

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

2. Lloyd regularly takes a break from work to go to the local cafe. The amount of time Lloyd waits to be served, in minutes, is modelled by the continuous random variable T, having probability density function

$$f(t) = \begin{cases} \frac{t}{120} & 4 \leqslant t \leqslant 16 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that the cumulative distribution function is given by

$$F(t) = \begin{cases} 0 & t < 4 \\ \frac{t^2}{240} - c & 4 \leqslant t \leqslant 16 \\ 1 & t > 16 \end{cases}$$

where the value of c is to be found.

(2)

(b) Find the exact probability that the amount of time Lloyd waits to be served is between 5 and 10 minutes.

(2)

(c) Find the median of T.

(2)

(d) Find the value of k such that

$$P(T < k) = \frac{2}{3} P(T > k)$$

giving your answer to 3 significant figures.

(3)

(a) For $4 \leqslant t \leqslant 16$: $F(t) = \int_4^t \frac{u}{120}\, du = \left[\frac{u^2}{240}\right]_4^t$

$= \frac{t^2}{240} - \frac{16}{240} = \frac{t^2}{240} - \frac{1}{15}$ so $c = \frac{1}{15}$

(b) $P(5 < T < 10) = F(10) - F(5) = \frac{100}{240} - \frac{1}{15} - \left(\frac{25}{240} - \frac{1}{15}\right)$

$= \frac{75}{240} = \frac{5}{16}$

(c) $F(m) = 0.5$: $\frac{m^2}{240} - \frac{1}{15} = \frac{1}{2} \Rightarrow \frac{m^2}{240} = \frac{17}{30} \Rightarrow m^2 = 136$

$m = \sqrt{136} = 2\sqrt{34} = 11.6619\ldots = 11.7$ minutes (3 s.f.)

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

stemup.co.uk

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

Question 2 continued

(d) Let $P(T < k) = p$, then $p = \frac{2}{3}(1 - p) \Rightarrow \frac{5}{3}p = \frac{2}{3} \Rightarrow p = \frac{2}{5}$

$F(k) = \frac{2}{5}$: $\frac{k^2}{240} - \frac{1}{15} = \frac{2}{5} \Rightarrow \frac{k^2}{240} = \frac{7}{15}$

$k^2 = 112 \Rightarrow k = \sqrt{112} = 10.5830...$

$k = 10.6$ (3 s.f.)

Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

Question 2 continued

(Total for Question 2 is 9 marks)

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

3. Two students, Jim and Dora, collected data on the mean annual rainfall, w cm, and the annual yield of leeks, l tonnes per hectare, for 10 years.

Jim summarised the data as follows

$$S_{wl} = 42.786 \qquad S_{ww} = 9936.9 \qquad \sum l^2 = 26.2326 \qquad \sum l = 16.06$$

(a) Find the product moment correlation coefficient between l and w

(2)

Dora decided to code the data first using $s = w - 6$ and $t = l - 20$

(b) Write down the value of the product moment correlation coefficient between s and t. Give a justification for your answer.

(1)

Dora calculates the equation of the regression line of t on s to be $t = 0.00431s - 18.87$

(c) Find the equation of the regression line of l on w in the form $l = a + bw$, giving the values of a and b to 3 significant figures.

(3)

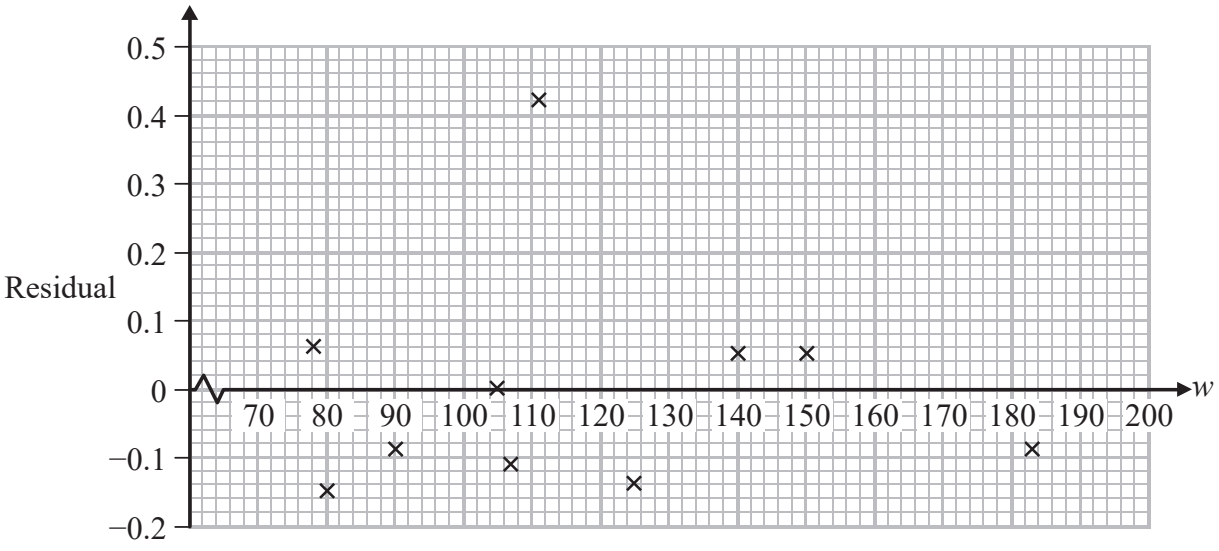
(d) Use your equation to estimate the yield of leeks when w is 100 cm.

(1)

(e) Calculate the residual sum of squares.

(2)

The graph shows the residual for each value of l



(f) (i) State whether this graph suggests that the use of a linear regression model is suitable for these data. Give a reason for your answer.

(ii) Other than collecting more data, suggest how to improve the fit of the model in part (c) to the data.

(2)

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

stemup.co.uk

Question 3 continued

(a) $S_{ll} = \Sigma l^2 - \frac{(\Sigma l)^2}{n} = 26.2326 - \frac{16.06^2}{10} = 0.44024$

$r = \frac{S_{wl}}{\sqrt{(S_{ww}S_{ll})}} = \frac{42.786}{\sqrt{(9936.9 \times 0.44024)}} = 0.646891...$

$r = 0.647$ (3 s.f.)

(b) $r = 0.647$, since coding (adding or subtracting constants) has no effect on the product moment correlation coefficient.

(c) $l - 20 = 0.00431(w - 6) - 18.87$

$l = 0.00431w - 0.02586 - 18.87 + 20$

$l = 0.00431w + 1.10414...$

$l = 1.10 + 0.00431w$ ($a = 1.10$, $b = 0.00431$)

(d) $l = 1.10 + 0.00431 \times 100 = 1.531$

Estimated yield = 1.53 tonnes per hectare

(e) $RSS = S_{ll} - \frac{(S_{wl})^2}{S_{ww}} = 0.44024 - \frac{42.786^2}{9936.9} = 0.44024 - 0.184227...$

$RSS = 0.256013... = 0.256$ (3 s.f.)

(f)(i) Yes, a linear regression model is suitable: the residuals appear to be randomly scattered above and below zero, with no obvious pattern.

(ii) There is a possible outlier (the point with residual about 0.42 at $w \approx 113$). It could be removed and the regression line recalculated.

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

stemup.co.uk

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 11 marks)

4. The random variable X has a continuous uniform distribution over the interval $[5, a]$, where a is a constant.

Given that $\text{Var}(X) = \frac{27}{4}$

(a) show that $a = 14$

(3)

The continuous random variable Y has probability density function

$$f(y) = \begin{cases} \frac{1}{20}(2y - 3) & 2 \leqslant y \leqslant 6 \\ 0 & \text{otherwise} \end{cases}$$

The random variable $T = 3(X^2 + X) + 2Y$

(b) Show that $E(T) = \frac{9857}{30}$

(7)

(a) $X \sim U[5, a]$: $\text{Var}(X) = \frac{(a-5)^2}{12} = \frac{27}{4}$

$(a-5)^2 = 81$

$a - 5 = 9$ or $a - 5 = -9 \Rightarrow a = 14$ or $a = -4$

Since $a > 5$, $a = -4$ is rejected, so $a = 14$ as required

(b) $E(Y) = \int_2^6 \frac{1}{20} y(2y-3)\,dy = \frac{1}{20}\int_2^6 (2y^2 - 3y)\,dy$

$= \frac{1}{20}\left[\frac{2y^3}{3} - \frac{3y^2}{2}\right]_2^6 = \frac{1}{20}\left((144 - 54) - \left(\frac{16}{3} - 6\right)\right)$

$= \frac{1}{20} \times \frac{272}{3} = \frac{68}{15}$

$E(X) = \frac{5 + 14}{2} = 9.5$

$\text{Var}(X) = E(X^2) - [E(X)]^2 \Rightarrow \frac{27}{4} = E(X^2) - 9.5^2$

$E(X^2) = 6.75 + 90.25 = 97$

$E(X^2 + X) = 97 + 9.5 = 106.5$

$E(T) = 3E(X^2) + 3E(X) + 2E(Y)$

$= 3 \times 97 + 3 \times 9.5 + 2 \times \frac{68}{15} = 319.5 + \frac{136}{15}$

$= \frac{9585}{30} + \frac{272}{30} = \frac{9857}{30}$ as required

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

stemup.co.uk

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

Question 4 continued

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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

Question 4 continued

Question 4 continued

(Total for Question 4 is 10 marks)

TOTAL FOR FURTHER STATISTICS 2 IS 40 MARKS

END

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk