Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number Candidate Number

Pearson Edexcel Level 3 GCE

Friday 19 May 2023

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

1. Every applicant for a job at *Donala* is given three different tasks, P, Q and R.
For each task the applicant is awarded a score.
The scores awarded to 9 of the applicants, for the tasks P and Q, are given below.

Applicant	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*
Task *P*	19 1	16 3	16 4	12 5	8 8	17 2	12 6	12 7	5 9
Task *Q*	17 2	11 5	14 4	7 8	6 9	18 1	15 3	11 6	10 7

(a) Calculate Spearman's rank correlation coefficient for the scores awarded for the tasks P and Q.

(4)

(b) Test, at the 1% level of significance, whether or not there is evidence for a positive correlation between the ranks of scores for tasks P and Q.
You should state your hypotheses and critical value clearly.

(4)

The Spearman's rank correlation coefficient for P and R is 0.290 and for Q and R is 0.795

The manager of *Donala* wishes to reduce the number of tasks given to job applicants from three to two.

(c) Giving a reason for your answer, state which 2 tasks you would recommend the manager uses.

(2)

(a) Let Rank 1 = 19 and rank all other data points accordingly:

Ranks	A	B	C	D	E	F	G	H	I	
P	1	3.5	3.5	6	8	2	6	6	9	1
Q	2	5.5	4	8	9	1	3	5.5	7	1

see table in Q1. 5th and 6th data point have the same value (11) so rank both as $\frac{5+6}{2} = 5.5$

$\Sigma p^2 = 1^2 + 2(3.5)^2 + 8^2 + 2^2 + 3(6)^2 + 9 = 282.5$

$\Sigma q^2 = 2^2 + 2(5.5)^2 + 8^2 + 9^2 + 1^2 + 3^2 + 7^2 = 284.5$

$\Sigma p = 45 = \Sigma q$

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

Question 1 continued

$\Sigma pq = (1\times2)+(3.5\times5.5)+(3.5\times4)+(6\times8)+(8\times9)+(2\times1)$

$+(6\times3)+(6\times5.5)+(9\times7)$

$\Sigma pq = 271.25$ 1

$$r_s = \frac{271.25 - \frac{45^2}{9}}{\sqrt{\left(282.5-\frac{45^2}{9}\right)\left(284.5-\frac{45^2}{9}\right)}} = 0.7907$$ 1

PMCC formula adapted for r_s using rankings.

(b) $H_0: \rho = 0$ and $H_1: \rho > 0$ ← one-tailed 1

Significance level = 1% = 0.01

Critical value: $\rho = 0.7833$ 1 ← using statistical table for Spearman's with $n=9$

$(r_s =)\ 0.7907 > 0.7833$ ∴ the result is significant and we reject H_0. 1

There is evidence of a positive correlation between the ranks of scores for tasks P and Q. 1

(c) The task pairs (P, Q) and (Q, R) have high coefficients, so provide similar information about the applicants. 1

Therefore the manager should keep tasks P and R 1

→ low correlation = provide different information

(Total for Question 1 is 10 marks)

2. A continuous random variable X has probability density function

$$f(x) = \begin{cases} \frac{x}{16}(9 - x^2) & 1 \leqslant x \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the cumulative distribution function of X **(3)**

(b) Calculate $P(X > 1.8)$ **(2)**

(c) Use calculus to find $E\left(\frac{3}{X} + 2\right)$ **(3)**

(d) Show that the mode of X is $\sqrt{3}$ **(3)**

(a) c.d.f $F(x) = P(X \leq x)$ which is the same as the area under the curve, so integrate:

$$f(x) = \frac{x}{16}(9 - x^2) = \frac{9x}{16} - \frac{x^3}{16}$$

$$\int \frac{9x}{16} - \frac{x^3}{16}\,dx = \frac{9x^2}{32} - \frac{x^4}{64} + c \quad 1$$

At $x = 1$, $\frac{9(1)^2}{32} - \frac{(1)^4}{64} = \frac{17}{64}$ ← since when $x < 1$, $f(x) = 0$, $F(x) = 0$ when $x = 1$ as it is cumulative

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{9x^2}{32} - \frac{x^4}{64} - \frac{17}{64} & 1 \leq x \leq 3 \quad 1 \\ 1 & x > 3 \quad 1 \end{cases}$$

cumulative probability so must end at 1.

DO NOT WRITE IN THIS AREA

Question 2 continued

(b) $P(X > 1.8) = 1 - F(1.8)$ ① $\leftarrow F(1.8) = P(X \leq 1.8)$

$$= 1 - \left[\frac{9(1.8)^2}{32} - \frac{(1.8)^4}{64} - \frac{17}{64}\right]$$

$$= 1 - 0.4816$$

$$= 0.5184$$ ①

(c) $E(X) = \int x f(x)\,dx$

$$E(X^{-1}) = \int_1^3 x^{-1}\left(\frac{9}{16}x - \frac{1}{16}x^3\right)dx$$ ①

$$E(X^{-1}) = \int_1^3 \left(\frac{9}{16} - \frac{1}{16}x^2\right)dx$$

$$E(X^{-1}) = \left[\frac{9}{16}x - \frac{1}{48}x^3\right]_1^3$$

$$E(3X^{-1}+2) = 3 \times \left[\frac{9}{16}x - \frac{1}{48}x^3\right]_1^3 + 2$$ ①

$$E(3X^{-1}+2) = 3 \times \left(\left[\frac{9}{16}(3) - \frac{1}{48}(3)^3\right] - \left[\frac{9}{16}(1) - \frac{1}{48}(1)^3\right]\right) + 2$$

$$E(3X^{-1}+2) = 3.75$$ ①

Question 2 continued

(d) $f(x) = \frac{9}{16}x - \frac{1}{16}x^3$

$f'(x) = \frac{9}{16} - \frac{3}{16}x^2$ 1 ← mode will be highest point on graph of p.d.f, so differentiate for turning point.

$\frac{9}{16} - \frac{3}{16}x^2 = 0$

$\frac{9}{16} = \frac{3}{16}x^2$ 1

$\times 16$

$9 = 3x^2$

$\div 3$

$3 = x^2$

$\pm\sqrt{3} = x$

Check location of boundaries.

$f(1) = \frac{9}{16}(1) - \frac{1}{16}(1)^3 = 0.5$

$f(1) = \frac{9}{16}(3) - \frac{1}{16}(3)^3 = 1$

$f(\sqrt{3}) = \frac{9}{16}(\sqrt{3}) - \frac{1}{16}(\sqrt{3})^3 = 0.6495$

$0.5 < 0.6495 < 1$ ∴ the mode lies within the boundary, so is a true mode 1

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

Question 2 continued

(Total for Question 2 is 11 marks)

3. Pat is investigating the relationship between the height of professional tennis players and the speed of their serve. Data from 9 randomly selected professional male tennis players were collected. The variables recorded were the height of each player, h metres, and the maximum speed of their serve, v km/h.

Pat summarised these data as follows

$$\sum h = 17.63 \qquad \sum v = 2174.9 \qquad \sum v^2 = 526\,407.8 \qquad S_{hh} = 0.0487 \qquad S_{hv} = 5.1376$$

(a) Calculate the product moment correlation coefficient between h and v **(2)**

(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 v on h in the form $v = a + bh$ where a and b are to be given to one decimal place. **(3)**

Pat calculated the sum of the residuals for the 9 tennis players as 1.04

(d) Without doing a calculation, explain how you know Pat has made a mistake. **(1)**

Pat made one mistake in the calculation. For the tennis player of height 1.96 m Pat misread the residual as 2.27

(e) Find the maximum speed of serve, in km/h, for the tennis player of height 1.96 m **(3)**

(a) $S_{vv} = \Sigma v^2 - \frac{(\Sigma v)^2}{9}$ $\qquad S_{xy} = \Sigma xy - \frac{(\Sigma xy)^2}{n}$

$S_{vv} = 526407.8 - \frac{2174.9^2}{9}$ $\qquad r = \frac{S_{xy}}{\sqrt{S_{xx}} \times \sqrt{S_{yy}}}$

$S_{vv} = 831.132$ ①

$r = \frac{5.1376}{\sqrt{831.132} \times \sqrt{0.0487}}$ ① $= 0.80753.$ ①

(b) There is a positive correlation which is close to 1, so it is consistent with a linear model. ①

DO NOT WRITE IN THIS AREA

Question 3 continued

(c) $V = a + bh$

intercept; b is the gradient

S_{hv} = Covariance(h,v)
$= \beta \times \text{Var}(h)$
$= \beta \times S_{hh}$
$\therefore \beta = \frac{S_{hv}}{S_{hh}}$

$b = \frac{S_{hv}}{S_{hh}} = \frac{5.1376}{0.0487} = 105.49$ (1)

$a = \frac{\Sigma v}{n} - b\left(\frac{\Sigma h}{n}\right)$ ← rearrange $v = a + bh$

$a = \frac{2174.9}{9} - 105.49 \times \left(\frac{17.63}{9}\right) = 35$ (1)

$V = 105.5h + 35.0$ (1) ← must be to 1dp.

(d) The sum of the residuals should be zero (1)

(e) Residual $= 2.27 - 1.04 = 1.23$ (1) ← calculate actual residual

$V = 105.5(1.96) + 35.0 + 1.23$ (1) ← from part (c)

$V = 243.01$ km/h (1)

(Total for Question 3 is 10 marks)

4. The random variable X has a continuous uniform distribution over the interval $[-3, k]$

Given that $\mathrm{P}(-4 < X < 2) = \frac{1}{3}$

(a) find the value of k

(3)

A computer generates a random number, Y, where

- Y has a continuous uniform distribution over the interval $[a, b]$
- $\mathrm{E}(Y) = 6$
- $\mathrm{Var}(Y) = 192$

The computer generates 5 random numbers.

(b) Calculate the probability that at least 2 of the 5 numbers generated are greater than 7.5

(6)

(a) $P(-4 < X < 2) = \frac{1}{3} = P(-3 < X < 2)$ ①

because -4 is not included in the interval.

$\frac{2--3}{k--3} = \frac{1}{3}$ ① ← $2--3$ is the width of the first interval.
← $k--3$ is the proportional width of the unknown interval.

$(2+3)\times 3 = k+3$

$15 = k+3$

$12 = k$ ①

(b) $E(Y) = 6 \Rightarrow \frac{a+b}{2} = 6 \Rightarrow a+b = 12$

$Var(Y) = 192 \Rightarrow \frac{1}{12}(b-a)^2 = 192$

using uniform distribution formulae 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

$a = 12 - b \Rightarrow \frac{1}{12}(b - 12 + b)^2 = 192$ ①

$2b - 12 = \sqrt{2304}$

$2b = 48 + 12$

$b = 30$ ①

$a = 12 - 30 \Rightarrow a = -18$

$P(Y > 7.5) = \frac{30 - 7.5}{30 - (-18)} = 0.46875$ ①

this is probability p

$R \sim B(5, 0.46875)$ ① ← apply binomial distribution

$P(R \geq 2) = 1 - P(R \leq 1)$

from stat tables.

$= 0.7710$ ①

Question 4 continued

(Total for Question 4 is 9 marks)

TOTAL FOR FURTHER STATISTICS 2 IS 40 MARKS