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 **8FM0/23**

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 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

Pearson

1. The discrete random variable X has the following distribution

x	0	1	2	3	4
$\mathrm{P}(X = x)$	r	k	$\frac{k}{2}$	$\frac{k}{3}$	$\frac{k}{4}$

where r and k are positive constants.

The standard deviation of X equals the mean of X

Find the exact value of r

(6)

Formulae for Mean and Variance

$E(X) = \sum x P(X = x)$ mean, μ

$E(X^2) = \sum x^2 P(X = x)$ mean of squares

$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$ variance, σ^2

standard deviation $\sigma = \sqrt{\mathrm{Var}(X)}$

x	0	1	2	3	4
x^2	0	1	4	9	16
$P(X=x)$	r	k	$\frac{k}{2}$	$\frac{k}{3}$	$\frac{k}{4}$

← this row for $E(X)$

← this row for $E(X^2)$

$E(X) = 0 \times r + 1 \times k + \cancel{2} \times \frac{k}{\cancel{2}} + \cancel{3} \times \frac{k}{\cancel{3}} + \cancel{4} \times \frac{k}{\cancel{4}}$

$= k + k + k + k$

$E(X) = 4k$ mean

$E(X^2) = 0 \times r + 1 \times k + \overset{2}{\cancel{4}} \times \frac{k}{\cancel{2}} + \overset{3}{\cancel{9}} \times \frac{k}{\cancel{3}} + \overset{4}{\cancel{16}} \times \frac{k}{\cancel{4}}$

$= k + 2k + 3k + 4k$

$E(X^2) = 10k$ M1B1

Substitute into formula for $\mathrm{Var}(X)$:

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

$= 10k - (4k)^2$

$= 10k - 16k^2 = \sigma^2$

$\sigma = \sqrt{10k - 16k^2}$ standard deviation

Given that mean = standard deviation. Use this to find k:

$4k = \sqrt{10k - 16k^2}$ M1

$(4k)^2 = 10k - 16k^2$

$16k^2 = 10k - 16k^2$

$32k^2 - 10k = 0$

Solve quadratic for k:

$k(32k - 10) = 0$

$k = 0 \quad k = \frac{10}{32} = \frac{5}{16} \rightarrow k = \frac{5}{16}$ A1

P 7 2 8 0 9 A 0 2 1 2

Question 1 continued

Use the fact that $\sum$probabilities $= 1$ to get the value of r:

$r + k + \frac{k}{2} + \frac{k}{3} + \frac{k}{4} = 1$ $\quad \sum$prob. $= 1$

$r + \frac{5}{16} + \frac{1}{2}\left(\frac{5}{16}\right) + \frac{1}{3}\left(\frac{5}{16}\right) + \frac{1}{4}\left(\frac{5}{16}\right) = 1$ $\quad$ substitute $k = \frac{5}{16}$

$r = 1 - \frac{5}{16}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right)$ $\quad$ solve for r

$r = 1 - \frac{5}{16}\left(\frac{25}{12}\right)$

$r = 1 - \frac{125}{192}$ $\quad$ M1

$r = \frac{67}{192}$ $\quad$ value of r $\quad$ A1

(Total for Question 1 is 6 marks)

2. A bag contains a large number of balls, all of the same size and weight. The balls are coloured Red, Blue or Yellow.

Jasmine asks each child in a group of 150 children to close their eyes, select a ball from the bag and show it to her. The child then replaces the ball and repeats the process a second time.

If both balls are the same colour the child receives a prize.

The results are given in the table below.

2nd colour \ 1st colour	Red	Blue	Yellow	Total
Red	31	11	18	60
Blue	8	10	9	27
Yellow	21	9	33	63
Total	60	30	60	150

Jasmine carries out a test, at the 5% level of significance, to see whether or not the colour of the 2nd ball is independent of the colour of the 1st ball.

(a) Calculate the expected frequencies for the cases where both balls are the same colour.
(2)

The test statistic Jasmine obtained was 12.712 to three decimal places.

(b) Use this value to complete the test, stating the critical value and conclusion clearly.
(3)

With reference to your calculations in part (a) and the nature of the experiment,

(c) give a plausible reason why Jasmine may have obtained her conclusion in part (b).
(1)

(a) For contingency tables we calculate the expected frequency like this:

			Total
	a	b	a + b
	c	d	c + d
Total	a + c	b + d	a+b+c+d

row total

column total

grand total

to get the E_i of this, we calculate

$E_i = \dfrac{\text{row total} \times \text{column total}}{\text{grand total}}$

$\therefore$ here it would be $E_i = \dfrac{(a+b)(a+c)}{a+b+c+d}$

3 Cases:

Red - Red

$E(RR) = \dfrac{60 \times 60}{150} = 24$

$E(RR) = 24$

Blue - Blue

$E(BB) = \dfrac{30 \times 27}{150}$

$E(BB) = 5.4$

Yellow - Yellow

$E(YY) = \dfrac{60 \times 63}{150}$

$E(YY) = 25.2$

M1A1

P 7 2 8 0 9 A 0 4 1 2

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

Question 2 continued

(b) The box with the lowest row and column totals is blue-blue. Since in (a) we calculated E(BB) = 5.4,
the lowest E_i is >5 and $\therefore$ we have no pooling

★ For contingency tables the formula for DoF is:

DoF = (#of rows - 1)(#of columns - 1)

Substitute:

$$\text{DoF} = (3-1)(3-1)$$
$$= 2 \times 2$$
$$= 4 \text{ dof.}$$ B1

Critical value from tables:

$\chi^2_4(5\%) = 9.488$

Compare with given test statistic:

$12.712 > 9.488 \therefore$ falls in critical region. B1

$\therefore$ sufficient evidence that the color of balls is not independent. B1

(c) Observed frequency is larger than expected frequency for balls of the same colour.
Some children may have cheated when choosing the 2^{nd} ball $\therefore$ choice not independent B1

(Total for Question 2 is 6 marks)

3. A machine produces cloth. Faults occur randomly in the cloth at a rate of 0.4 per square metre.

The machine is used to produce tablecloths, each of area A square metres. One of these tablecloths is taken at random.

The probability that this tablecloth has no faults is 0.0907

(a) Find the value of A

(3)

The tablecloths are sold in packets of 20

A randomly selected packet is taken.

(b) Find the probability that more than 1 of the tablecloths in this packet has no faults.

(3)

A hotel places an order for 100 tablecloths each of area A square metres.

The random variable X represents the number of these tablecloths that have no faults.

(c) Find

(i) $\mathrm{E}(X)$

(ii) $\mathrm{Var}(X)$

(3)

(d) Use a Poisson approximation to estimate $\mathrm{P}(X = 10)$

(2)

It is claimed that a new machine produces cloth with a rate of faults that is less than 0.4 per square metre.

A piece of cloth produced by this new machine is taken at random.
The piece of cloth has area 30 square metres and is found to have 6 faults.

(e) Stating your hypotheses clearly, use a suitable test to assess the claim made for the new machine. Use a 5% level of significance.

(4)

(f) Write down the p-value for the test used in part (e).

(1)

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

Question 3 continued

(a) We're given a rate ∴ we are using Poisson Distribution.

$X \to$ # of faults per A square meters   define variable

$X \sim Po(\underbrace{0.4}_{\text{value per 1 square meter}} \times A)$ M1

We are also given that

$P(X=0)=0.0907$

Substitute: $x=0$, $\lambda=0.4A$ and $P(X=x)=0.0907$:   Formula for Poisson Distribution: $P(X=x)=\frac{e^{-\lambda}\times\lambda^{x}}{x!}$

$$0.0907=\frac{e^{-0.4A}\times \cancel{(0.4A)^{0}}}{\cancel{0!}}$$ M1

$0.0907=e^{-0.4A}$ apply ln on both sides

$\ln 0.0907=\ln e^{-0.4A}$ ★ ln rule: $\ln e^{x}=x$

$\ln 0.0907=-0.4A$

$A=\frac{\ln 0.0907}{-0.4}=6$

$A=6$ value of A   A1

(b) $T \to$ # of packets with no faults   define variable

$T \sim B(20, 0.0907)$ M1

$P(T>1)=1-P(T\leq 1)$ M1

$=0.55276 \longrightarrow 0.553$ to 3sf   A1

(c) $X \sim$ # of table cloths with no faults   define variable

$X \sim B(100, 0.0907)$

For Binomial:

$E(X)=np$

$Var(X)=np(1-p)$

i. $np=100\times 0.0907$

$E(X)=9.07$ M1A1

ii. $np(1-p)=9.07(1-0.0907)$

$Var(X)=8.247351 \longrightarrow Var(X)=8.25$ to 3sf   A1

(d) We know that λ is the mean, $\therefore \lambda=9.07$ (from (c)i.)

$X \sim\approx Po(9.07)$ M1

$P(X=10)=0.11947 \longrightarrow 0.119$ to 3sf   A1

Question 3 continued

(e) Hypotheses

$H_0: \lambda = 0.4$ $\xrightarrow{\times 30}$ $H_0: \lambda = 12$

$H_1: \lambda < 0.4$ $H_1: \lambda < 12$ B1

for 1 square meter for 30 square meters

We are given the found value 6.

$Y \rightarrow$ # of faults made in $30m^2$ from new machine

$Y \sim Po(12)$ M1

$P(Y \leq 6) = 0.04582 < 0.05$ $\therefore$ 6 falls in critical region A1

Significant evidence to reject H_0. The claim is supported A1

(f) p-value is the actual significance

$\therefore$ p-value $\rightarrow$ 0.0458 B1

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 16 marks)

anagram of "Pearson" :P

4. Table 1 below shows the number of car breakdowns in the *Snoreap* district in each of 60 months.

Number of car breakdowns	0	1	2	3	4	5
Frequency	12	11	19	14	3	1

Table 1

Anja believes that the number of car breakdowns per month in *Snoreap* can be modelled by a Poisson distribution. Table 2 below shows the results of some of her calculations.

Number of car breakdowns	0	1	2	3	4	$\geqslant 5$
Observed frequency (O_i)	12	11	19	14	3	1
Expected frequency (E_i)	9.92			9.64	4.34	

Table 2

(a) State suitable hypotheses for a test to investigate Anja's belief. **(1)**

(b) Explain why Anja has changed the label of the final column to $\geqslant 5$ **(1)**

(c) Showing your working clearly, complete Table 2 **(4)**

(d) Find the value of $\frac{(O_i - E_i)^2}{E_i}$ when the number of car breakdowns is

(i) 1

(ii) 3 **(2)**

(e) Explain why Anja used 3 degrees of freedom for her test. **(2)**

The test statistic for Anja's test is 6.54 to 2 decimal places.

(f) Stating the critical value and using a 5% level of significance, complete Anja's test. **(2)**

(a) Hypotheses

H_0: The number of breakdowns per month can be modelled by a Poisson Distribution

H_1: The number of breakdowns per month can't be modelled by a Poisson Distribution B1

(b) Poisson distribution would give all values above 5 a probability since theoretically the values are infinite B1

P 7 2 8 0 9 A 0 1 0 1 2

Question 4 continued

(c) We need λ. Used observed frequencies to calculate:

$\lambda = \frac{0\times12+1\times11+2\times19+3\times14+4\times3+5\times1}{12+11+19+14+3+1} = 1.8 \rightarrow \lambda=1.8$ for 1 month. M1

$X \rightarrow$ # of breakdowns in 1 month

Using $X \sim Po(1.8)$:

$E_1 = 60 \times P(X=1) = 17.85$ to 2dp M1 A1

$E_2 = 60 \times P(X=2) = 16.06$ to 2dp

$E_5 = 60 - E_{0-4} = 60 - 57.81 \rightarrow (9.92+17.85+16.06+9.64+4.34 = 57.81)$

$= 2.18$ to 2dp B1

(d) i. $\frac{(11-17.85)^2}{17.85} = 2.6287... \rightarrow 2.63$ to 3sf.

ii. $\frac{(14-9.64)^2}{9.64} = 1.97195... \rightarrow 1.97$ to 3sf M1 A1

(e) We don't need to combine the last two columns since $E_i > 5$. B1

We calculated λ above using O_i $\therefore$ subtract 2 degrees of freedom: B1

$5-2=3$ DoF

(f) Get critical value from tables

B1 $\chi^2_3(5\%) = 7.815 > 6.54$ $\therefore$ doesn't fall in critical region.

Insufficient evidence to reject H_0.

Insufficient evidence to reject her belief. B1

Question 4 continued

(Total for Question 4 is 12 marks)

TOTAL FOR FURTHER STATISTICS 1 IS 40 MARKS