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

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 the statistical tables should be quoted in full. If a calculator is used instead of the 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.
- Good luck with your examination.

Turn over ▶

P66791RRA

©2021 Pearson Education Ltd.

1/1/1/1/1/1/1/1/

P 6 6 7 9 1 R R A 0 1 1 6

Pearson

1. *Flobee* sells tomato seeds in packets, each containing 40 seeds. *Flobee* advertises that only 4% of its tomato seeds do not germinate.

Amodita is investigating the germination of *Flobee*'s tomato seeds. She plants 125 packets of *Flobee*'s tomato seeds and records the number of seeds that do not germinate in each packet.

Number of seeds that do not germinate	0	1	2	3	4	5	6 or more
Frequency	15	35	38	22	10	5	0

Amodita wants to test whether the binomial distribution B(40, 0.04) is a suitable model for these data.

The table below shows the expected frequencies, to 2 decimal places, using this model.

Number of seeds that do not germinate	0	1	2	3	4	5 or more
Expected Frequency	24.42	40.70	r	17.45	6.73	s

(a) Calculate the value of r and the value of s

(2)

(b) Stating your hypotheses clearly, carry out the test at the 5% level of significance. You should state the number of degrees of freedom, critical value and conclusion clearly.

(6)

Amodita believes that *Flobee* should use a more realistic value for the percentage of their tomato seeds that do not germinate.
She decides to test the data using a new model B(40, p)

(c) Showing your working, suggest a more realistic value for p

(2)

(a) $X \sim B(40, 0.04)$

$r = 125 \times P(X = 2) = 125 \times 0.264558... = 33.0697...$

$s = 125 \times [1 - P(X \leq 4)] = 125 \times 0.0210223... = 2.6277...$

$r = 33.07, \quad s = 2.63$

(b) H_0: B(40, 0.04) is a suitable model for the data

H_1: B(40, 0.04) is not a suitable model for the data

Expected frequencies for 4 and 5 or more are < 5, so combine the last two cells into "4 or more":

$O = 10 + 5 = 15, \quad E = 6.73 + 2.63 = 9.36$

Number	0	1	2	3	4 or more
O	15	35	38	22	15
E	24.42	40.70	33.07	17.45	9.36

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

Question 1 continued

$(O - E)^2 \div E$ 3.6338 0.7983 0.7350 1.1864 3.3985

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} = \frac{(15 - 24.42)^2}{24.42} + \frac{(35 - 40.7)^2}{40.7} + \frac{(38 - 33.07)^2}{33.07} + \ldots$$

$$+ \frac{(22 - 17.45)^2}{17.45} + \frac{(15 - 9.36)^2}{9.36}$$

$\chi^2 = 3.6338 + 0.7983 + 0.7350 + 1.1864 + 3.3985 = 9.752$

Degrees of freedom: 5 cells, 1 constraint $\Rightarrow \nu = 5 - 1 = 4$

Critical value: $\chi^2_4(0.05) = 9.488$

$9.752 > 9.488$, so the result is significant: reject H_0

There is evidence that B(40, 0.04) is not a suitable model, so

Flobee's claim of 4% not germinating is not supported.

(c) Total not germinating $= 0\times15 + 1\times35 + 2\times38 + 3\times22 + 4\times10 + 5\times5$

$= 0 + 35 + 76 + 66 + 40 + 25 = 242$

$$p = \frac{242}{125 \times 40} = \frac{242}{5000} = 0.0484$$

A more realistic value is $p = 0.0484$

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

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

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. Rowan and Alex are both check-in assistants for the same airline.
The number of passengers, R, checked in by Rowan during a 30-minute period can be modelled by a Poisson distribution with mean 28

(a) Calculate $P(R \geqslant 23)$

(1)

The number of passengers, A, checked in by Alex during a 30-minute period can be modelled by a Poisson distribution with mean 16, where R and A are independent.
A randomly selected 30-minute period is chosen.

(b) Calculate the probability that exactly 42 passengers in total are checked in by Rowan and Alex.

(2)

The company manager is investigating the rate at which passengers are checked in.
He randomly selects 150 non-overlapping 60-minute periods and records the total number of passengers checked in by Rowan and Alex, in each of these 60-minute periods.

(c) Using a Poisson approximation, find the probability that for at least 25 of these 60-minute periods Rowan and Alex check in a total of fewer than 80 passengers.

(4)

On a particular day, Alex complains to the manager that the check-in system is working slower than normal. To see if the complaint is valid the manager takes a random 90-minute period and finds that the total number of people **Rowan** checks in is 67

(d) Test, at the 5% level of significance, whether or not there is evidence that the system is working slower than normal. You should state your hypotheses and conclusion clearly and show your working.

(4)

(a) $R \sim Po(28)$

$P(R \geqslant 23) = 1 - P(R \leqslant 22) = 1 - 0.14829... = 0.85170...$

$P(R \geqslant 23) = 0.852$

(b) $R \sim Po(28)$, $A \sim Po(16)$ independent

$Y = R + A \sim Po(28 + 16) = Po(44)$

$P(Y = 42) = \dfrac{e^{-44} \times 44^{42}}{42!} = 0.058664... = 0.0587$

(c) In 60 minutes the total $G \sim Po(2 \times 44) = Po(88)$

$P(G < 80) = P(G \leqslant 79) = 0.18320...$

Let X = number of the 150 periods with fewer than 80 passengers

$X \sim B(150, 0.18320...)$

Mean = $np = 150 \times 0.18320... = 27.48...$

Approximate by $T \sim Po(27.48...)$

$P(X \geqslant 25) \approx P(T \geqslant 25) = 1 - P(T \leqslant 24)$

$= 1 - 0.29224... = 0.70775...$

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

Question 2 continued

P(at least 25 periods with fewer than 80) = 0.708

(d) Rowan in 90 minutes: $J \sim Po(3 \times 28) = Po(84)$

H_0: $\lambda = 84$ H_1: $\lambda < 84$

(λ = mean number checked in by Rowan in 90 minutes)

Under H_0, $J \sim Po(84)$

$P(J \leq 67) = 0.03246...$

$0.0325 < 0.05$, so the result is significant: reject H_0

(67 lies in the critical region $J \leq 68$)

There is evidence at the 5% level of significance that the check-in system is working slower than normal, so the evidence supports Alex's complaint.

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

Question 2 continued

(Total for Question 2 is 11 marks)

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

3. The discrete random variable X has probability distribution

x	-3	-2	-1	0	2	5
$P(X = x)$	0.3	0.15	0.1	0.15	0.1	0.2

(a) Find $E(X)$

(1)

Given that $\text{Var}(X) = 8.79$

(b) find $E(X^2)$

(2)

The discrete random variable Y has probability distribution

y	-2	-1	0	1	2
$P(Y = y)$	$3a$	a	b	a	c

where a, b and c are constants.

For the random variable Y

$$P(Y \leqslant 0) = 0.75 \quad \text{and} \quad E(Y^2 + 3) = 5$$

(c) Find the value of a, the value of b and the value of c

(5)

The random variable $W = Y - X$ where Y and X are independent.

The random variable $T = 3W - 8$

(d) Calculate $P(W > T)$

(4)

(a) $E(X) = -3(0.3) - 2(0.15) - 1(0.1) + 0(0.15) + 2(0.1) + 5(0.2)$

$= -0.9 - 0.3 - 0.1 + 0 + 0.2 + 1 = -0.1$

$E(X) = -0.1$

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

$8.79 = E(X^2) - (-0.1)^2$

$E(X^2) = 8.79 + 0.01 = 8.8$

(c) $E(Y^2 + 3) = 5 \Rightarrow E(Y^2) = 2$

$(-2)^2 \times 3a + (-1)^2 \times a + 0^2 \times b + 1^2 \times a + 2^2 \times c = 2$

$12a + a + a + 4c = 2 \Rightarrow 14a + 4c = 2 \Rightarrow 7a + 2c = 1$ ①

$P(Y \leq 0) = 3a + a + b = 0.75 \Rightarrow 4a + b = 0.75$ ②

$\Sigma P(Y = y) = 1$: $3a + a + b + a + c = 1 \Rightarrow 5a + b + c = 1$ ③

③ − ②: $a + c = 0.25 \Rightarrow c = 0.25 - a$

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 3 continued

Sub into ①: $7a + 2(0.25 - a) = 1 \Rightarrow 5a + 0.5 = 1 \Rightarrow a = 0.1$

$c = 0.25 - 0.1 = 0.15, \quad b = 0.75 - 4(0.1) = 0.35$

$a = 0.1, \quad b = 0.35, \quad c = 0.15$

(d) $P(W > T) = P(W > 3W - 8) = P(8 > 2W) = P(W < 4)$

$W = Y - X \geq 4$ only for $(X, Y) = (-3, 1), (-3, 2), (-2, 2)$

$P(W \geq 4) = P(X = -3)P(Y = 1) + P(X = -3)P(Y = 2)$

$+ P(X = -2)P(Y = 2)$

$= 0.3 \times 0.1 + 0.3 \times 0.15 + 0.15 \times 0.15$

$= 0.03 + 0.045 + 0.0225 = 0.0975$

$P(W > T) = P(W < 4) = 1 - 0.0975 = 0.9025$

Question 3 continued

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 3 continued

(Total for Question 3 is 12 marks)

4. Charlie carried out a survey on the main type of investment people have.
The contingency table below shows the results of a survey of a random sample of people.

		Main type of investment		
		Bonds	Cash	Stocks
Age	25–44	a	$b - e$	e
	45–75	c	$d - 59$	59

(a) Find an expression, in terms of a, b, c and d, for the difference between the observed and the expected value $(O - E)$ for the group whose main type of investment is Bonds and are aged 45–75
Express your answer as a single fraction in its simplest form.

(4)

Given that $\sum \frac{(O-E)^2}{E} = 9.62$ for this information,

(b) test, at the 5% level of significance, whether or not there is evidence of an association between the age of a person and the main type of investment they have.
You should state your hypotheses, critical value and conclusion clearly.
You may assume that no cells need to be combined.

(3)

(a) Row total (45–75) $= c + (d - 59) + 59 = c + d$

Column total (Bonds) $= a + c$, Grand total $= a + b + c + d$

$$E = \frac{(c+d)(a+c)}{a+b+c+d}$$

$$O - E = c - \frac{(c+d)(a+c)}{a+b+c+d} = \frac{c(a+b+c+d) - (c+d)(a+c)}{a+b+c+d}$$

$$= \frac{ac + bc + c^2 + cd - ac - c^2 - ad - cd}{a+b+c+d}$$

$$O - E = \frac{bc - ad}{a+b+c+d}$$

(b) H_0: there is no association between the age of a person and the main type of investment they have

H_1: there is an association between the age of a person and the main type of investment they have

$\nu = (2-1)(3-1) = 2$, critical value $\chi^2_2(0.05) = 5.991$

$9.62 > 5.991$, so the result is significant: reject H_0

There is evidence of an association between the age of a person and the main type of investment they have.

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

Question 4 continued

(Total for Question 4 is 7 marks)

TOTAL FOR FURTHER STATISTICS 1 IS 40 MARKS

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