Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | Other names

Centre Number | Candidate Number

Pearson Edexcel Level 3 GCE

Friday 16 May 2025

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 ▶

P76372A

©2025 Pearson Education Ltd.
Y:1/1/1/

P 7 6 3 7 2 A 0 1 1 6

Pearson

1. A researcher is investigating the relationship between a person's age and their preferred method of shopping.

A random sample of 300 people is taken and the results are summarised in the table below.

		Preferred method of shopping		
		Online	**In-store**	**Total**
Age	**18 – 30**	27	23	50
	31 – 50	31	59	90
	51 – 64	17	43	60
	65 and older	30	70	100
	Total	105	195	300

(a) Write down suitable hypotheses for a test to determine whether or not there is evidence of an association between age and preferred method of shopping.

(1)

The researcher assumes that the null hypothesis is true and uses the data in the table to find expected values.

(b) (i) Identify the cell which would have the smallest expected value.

(ii) Calculate the expected value for this cell.

(2)

The value of $\sum \frac{(O-E)^2}{E}$ for the other 7 cells is 5.060

(c) Test, at the 5% level of significance, whether or not there is evidence of an association between age and preferred method of shopping.
You should state the test statistic, degrees of freedom, critical value and conclusion clearly.

(5)

(d) Explain whether or not your conclusion to part (c) would be the same if the test was carried out at a 1% level of significance.

(2)

(a) H_0: There is no association between age and preferred method of shopping

H_1: There is an association between age and preferred method of shopping

(b) (i) Online, aged 18 – 30 (the cell with observed frequency 27)

(ii) $E = \frac{50 \times 105}{300} = 17.5$

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 1 continued

(c) $\Sigma\frac{(O-E)^2}{E} = \frac{(27-17.5)^2}{17.5} + 5.060$

$= 5.15714... + 5.060 = 10.217... = 10.2$

Degrees of freedom $\nu = (4-1)(2-1) = 3$

Critical value: $\chi^2_3(0.05) = 7.815$

$10.2 > 7.815$, so the result is significant: reject H_0

There is evidence of an association between age and preferred method of shopping.

(d) At the 1% level the critical value is $\chi^2_3(0.01) = 11.345$

$10.2 < 11.345$, so the result would not be significant: H_0 would not be rejected. The conclusion would not be the same (it is reversed): there would be no evidence of an association at the 1% level.

Question 1 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 1 continued

(Total for Question 1 is 10 marks)

2. The discrete random variable X represents the score when a spinner is spun.
The probability distribution of X is given by

x	2	5	9
$P(X = x)$	0.6	0.3	0.1

(a) Find Var(X)
Show your working clearly.

(4)

A game is played by spinning the spinner twice.

If the two scores are the same, the number of **points** earned is 0

If the two scores are different, the number of **points** earned is the sum of the two scores.

(b) Show that the probability of earning 14 **points** in one game is 0.06

(1)

(c) Find the expected number of **points** earned when the game is played once.

(4)

Mehmet plays the game 150 times.

(d) Using a Poisson approximation, find the probability that Mehmet earns 14 **points** in exactly 4 of the games.

(3)

(a) $E(X) = 2 \times 0.6 + 5 \times 0.3 + 9 \times 0.1 = 1.2 + 1.5 + 0.9 = 3.6$

$E(X^2) = 2^2 \times 0.6 + 5^2 \times 0.3 + 9^2 \times 0.1 = 2.4 + 7.5 + 8.1 = 18$

$Var(X) = E(X^2) - [E(X)]^2 = 18 - 3.6^2 = 18 - 12.96$

$Var(X) = 5.04$

(b) 14 points needs one score of 5 and one score of 9 (either order)

$P(14 \text{ points}) = 2 \times 0.3 \times 0.1 = 0.06$

(c) Let T = points earned. Possible values of T: 0, 7, 11, 14

$P(T = 0) = 0.6^2 + 0.3^2 + 0.1^2 = 0.46$

$P(T = 7) = 2 \times 0.6 \times 0.3 = 0.36$ (scores 2 and 5)

$P(T = 11) = 2 \times 0.6 \times 0.1 = 0.12$ (scores 2 and 9)

$P(T = 14) = 2 \times 0.3 \times 0.1 = 0.06$ (scores 5 and 9)

$E(T) = 0 \times 0.46 + 7 \times 0.36 + 11 \times 0.12 + 14 \times 0.06$

$= 2.52 + 1.32 + 0.84$

$E(T) = 4.68$

(d) Y = number of games in which 14 points are earned

$Y \sim B(150, 0.06)$

n is large and p is small, so $Y \approx Po(150 \times 0.06) = Po(9)$

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 2 continued

$P(Y = 4) \approx \frac{e^{-9} \times 9^4}{4!} = 0.033737...$

$P(Y = 4) = 0.0337$

Question 2 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 2 continued

(Total for Question 2 is 12 marks)

3. Raoul, Steffi and Taro are catching butterflies for research.

The number of butterflies caught by Raoul per hour may be assumed to follow a Poisson distribution with mean 4.2

Find the probability that Raoul catches

(a) (i) exactly 6 butterflies in a randomly selected one-hour period,

(1)

(ii) exactly 1 butterfly in a randomly selected 10-minute period.

(2)

Following a long period without rain, Raoul believes there will now be a change to the rate at which he catches butterflies.
To test his belief, he uses the random variable R to represent the number of butterflies he catches in a 4-hour period.

A hypothesis test is to be carried out to determine whether or not there is support for Raoul's belief. The null hypothesis of the test will be rejected if $R \leqslant 9$ or $R > 25$

(b) Stating the hypotheses clearly, find the actual level of significance of the test.

(4)

The number of butterflies caught by Steffi per hour may be assumed to follow a Poisson distribution with mean 3.2

(c) Find the probability that in exactly 2 of the next 4 hours, Steffi catches less than or equal to 3 butterflies each hour.

(3)

The number of butterflies caught by Taro per hour may be assumed to follow a Poisson distribution with mean 2.7

Steffi and Taro both go to catch butterflies in a field one day.

Taro models the **total** number of butterflies caught per hour with a Poisson distribution with mean 3.2 + 2.7 = 5.9

(d) State a condition that would be needed for Taro's model to be valid.

(1)

(a) X = number caught by Raoul in one hour, $X \sim Po(4.2)$

(i) $P(X = 6) = \dfrac{e^{-4.2} \times 4.2^6}{6!} = 0.11432... = 0.114$

(ii) Y = number caught in 10 minutes, $Y \sim Po(4.2 \div 6) = Po(0.7)$

$P(Y = 1) = 0.7e^{-0.7} = 0.34760... = 0.348$

(b) λ = mean number of butterflies caught per hour

$H_0: \lambda = 4.2$ $H_1: \lambda \neq 4.2$

Under H_0: $R \sim Po(4 \times 4.2) = Po(16.8)$

$P(R \leqslant 9) = 0.02896...$

$P(R > 25) = 1 - P(R \leqslant 25) = 1 - 0.97769... = 0.02230...$

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

Actual significance level = 0.02896... + 0.02230... = 0.05126...

Actual level of significance = 0.0513 (5.13%)

(c) S = number caught by Steffi in one hour, $S \sim Po(3.2)$

$P(S \leqslant 3) = 0.60251...$

J = number of hours (out of 4) with $S \leqslant 3$, $J \sim B(4, 0.60251...)$

$P(J = 2) = 6 \times (0.60251...)^2 \times (0.39748...)^2$

$= 0.34413... = 0.344$

(d) Steffi and Taro must catch butterflies independently of each other (the number Steffi catches does not affect the number Taro catches).

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 11 marks)

4. A 5-question test is taken by 80 students.

Each question was answered either correctly or incorrectly.

Number of questions answered correctly	0	1	2	3	4	5
Observed frequency	15	23	31	9	2	0

(a) Show that the proportion of questions answered correctly by these 80 students is 0.3

(2)

Anisa believes that the number of questions answered correctly can be modelled using a binomial distribution.
In order to test her belief, Anisa first calculates some of the expected frequencies.

Number of questions answered correctly	0	1	2	3	4	5
Expected frequency	13.45	r	s	10.58	2.27	0.19

(b) Find the value of r and the value of s

(3)

(c) Explain fully why there are 2 degrees of freedom for the test of Anisa's belief.

(2)

(a) Mean $= \frac{0 \times 15 + 1 \times 23 + 2 \times 31 + 3 \times 9 + 4 \times 2 + 5 \times 0}{80} = \frac{120}{80} = 1.5$

Proportion $p = \frac{1.5}{5} = 0.3$ as required

(b) $X \sim B(5, 0.3)$

$P(X = 1) = 5 \times 0.3 \times 0.7^4 = 0.36015$

$r = 80 \times 0.36015 = 28.812 = 28.8$

$P(X = 2) = 10 \times 0.3^2 \times 0.7^3 = 0.3087$

$s = 80 \times 0.3087 = 24.696 = 24.7$

(c) The expected frequencies for 4 and 5 correct (2.27 and 0.19) are less than 5, and 2.27 + 0.19 = 2.46 is still less than 5, so the last 3 columns must be pooled (10.58 + 2.27 + 0.19 = 13.04 > 5), leaving 4 cells.
There are 2 constraints: the total frequency of 80, and the proportion $p = 0.3$ was estimated from the observed data.
Degrees of freedom = 4 − 2 = 2

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