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 17 May 2024

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

P75673A

F:1/1/1/

P 7 5 6 7 3 A 0 1 1 6

1. Sharma believes that each computer game he sells appeals equally to all age ranges.
To investigate this, he takes a random sample of 100 people who play these games and asks them which of the games A, B or C they prefer.
The results are summarised in the table below.

Computer game		A	B	C
Age range	< 20	8	15	6
	$20 - 30$	21	12	9
	> 30	6	10	13

(a) Write down hypotheses for a suitable test to assess Sharma's belief.
(1)

(b) For the test, calculate the expected frequency for

(i) those players aged under 20 who prefer game C

(ii) those players aged between 20 and 30 who prefer game A
(2)

(c) State the degrees of freedom of the test statistic for this test.
(1)

Sharma correctly calculates the test statistic for this test to be 11.542 (to 3 decimal places).

(d) Using a 5% significance level, and stating your critical value, comment on Sharma's belief.
(2)

(a) H_0: There is no association between age range and preferred game

H_1: There is an association between age range and preferred game

B1

P 7 5 6 7 3 A 0 2 1 6

Question 1 continued

(b) 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

column total → a + c; row total → a + b; grand total → a+b+c+d

to get the E_i of this, we calculate

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

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

(i)

age \ game	A	B	C	
< 20	8	15	6	29
20 – 30	21	12	9	42
> 30	6	10	13	29
	35	37	28	100

$E_i = \frac{\text{row} \times \text{column}}{\text{grand}}$

$= \frac{29 \times 28}{100}$

$= 8.12$

(ii)

age \ game	A	B	C	
< 20	8	15	6	29
20 – 30	21	12	9	42
> 30	6	10	13	29
	35	37	27	100

$E_i = \frac{\text{row} \times \text{column}}{\text{grand}}$

$= \frac{42 \times 35}{100}$

$= 14.7$

M1A1

(c) We can figure out that there is no pooling by looking at the expected frequency of the cell with the lowest observed frequency. In this case that is game C & < 20, with an O_i of 6.

We calculated this E_i in (b)i. $E_i = 8.12 > 5$ ∴ no pooling.

★ Since the smallest O_i has no pooling, we can assume the other cells won't need either.

★ For contingency tables the formula for DoF is:

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

In this case we have 3 rows & 3 columns ∴

$(3-1) \times (3-1)$

$= 2 \times 2$

$= 4$ DoF B1

Question 1 continued

(d) We are given the test statistic, $\chi^2 = 11.542$

Get the c.v. from tables in the formula booklet

$\chi^2_4(0.05) = 9.488$ B1

4 DoF, found in (c); 5% significance level

We compare the test statistic to the c.v. :

$11.542 > 9.488$

$\therefore$ test statistic falls in the critical region

Since the test is significant we reject H_0.

His belief is not supported. There is evidence of an association between age range and computer game preference. B1

☆ Remember your hypotheses from (a)

H_0: There is no association between age range and preferred game

H_1: There is an association between age range and preferred game

Question 1 continued

(Total for Question 1 is 6 marks)

2. A manager keeps a record of accidents in a canteen.

Accidents occur randomly with an average of 2.7 per month. The manager decides to model the number of accidents with a Poisson distribution.

(a) Give a reason why a Poisson distribution could be a suitable model in this situation. **(1)**

(b) Assuming that a Poisson model is suitable, find the probability of

(i) at least 3 accidents in the next month, 1 month **(1)**

(ii) no more than 10 accidents in a 3-month period, 3 months **(2)**

(iii) at least 2 months with no accidents in an 8-month period. 8 months **(4)**

One day, two members of staff bump into each other in the canteen and each report the accident to the manager. The canteen manager is unsure whether to record this as one or two accidents.

Given that the manager still wants to model the number of accidents per month with a Poisson distribution,

(c) state

- a property of the Poisson distribution that the manager should consider when deciding how to record this situation
- whether the manager should record this as one or two accidents

(1)

The manager introduces some new procedures to try and reduce the average number of accidents per month.

During the following 12 months the total number of accidents is 22
The manager claims that the accident rate has been reduced.

(d) Use a 5% level of significance to carry out a suitable test to assess the manager's claim.
You should state your hypotheses clearly and the p-value used in your test. **(4)**

(a) Since: ★Remember conditions for poisson

→ accidents occur independantly

→ accidents occur at a constant rate

B1

Question 2 continued

(b) i. A → # of accidents in 1 month define variable

A ~ Po(2.7) given rate for 1 month

P(A ≥ 3) = 1 − P(A ≤ 2) 3−1, as we want P(A = 3) to be included

= 1 − 0.49362

= 0.50637 → 0.506 to 3sf B1

ii. B → # of accidents in 3 months define variable

B ~ Po(3 × 2.7) # of accidents in 1 month. to get λ for 3 months

∴ B ~ Po(8.1) we multiply by 3, assuming the rate stays the same

P(B ≤ 10) = 0.805837

↳ 0.806 to 3sf M1A1

iii. "at least 2 months with no accidents in an 8 month period"
binomial poisson binomial

→ we will use both Poisson and Binomial here.

The information we will use with poisson is "no accidents".

from (b)i.:

A → # of accidents in 1 month we are considering just 1 of the 8 months here.

A ~ Po(2.7) given rate for 1 month

$P(A=0) = \frac{e^{-2.7} \times 2.7^0}{0!}$ → Formula for Poisson: $P(X=0) = \frac{e^{-\lambda}\lambda^x}{x!}$

$= e^{-2.7}$

what we got is the probability of a month having no accidents, which will be our p in the binomial distribution

For binomial we know n = 8 ("8 months") and that $p = e^{-2.7}$.

∴ X → # of months with no accidents

$X \sim B(8, e^{-2.7})$ M1 A1

P(X ≥ 2) = 1 − P(X ≤ 1) M1

= 1 − 0.903542

= 0.096457 → 0.0965 to 3sf A1

(c) For Poisson, events must be independent ∴ he should record it as 1 accident. B1

Question 2 continued

(d) Hypotheses

$H_0: \lambda = 2.7$

$H_1: \lambda < 2.7$

We are considering 12 months, $\therefore$ $2.7 \times 12 = \mu = 32.4$ — rate of 1 month; rate for 12 months

$\therefore$ New hypotheses:

$H_0: \mu = 32.4$

$H_1: \mu < 32.4$ B1

$Y \rightarrow$ # of accidents per Year — define variable (12 months $\therefore$ use μ)

$Y \sim Po(32.4)$ M1

$P(Y \leq 22) = 0.03512 < 0.05$ — since the probability we found was smaller than 0.05, 22 falls in the critical region.

found value was given A1

there is sufficient evidence to reject H_0 and the manager's claim is supported. (# of accidents per month decreased)

A1

DO NOT WRITE IN THIS AREA

P 7 5 6 7 3 A 0 8 1 6

Question 2 continued

(Total for Question 2 is 13 marks)

3. The discrete random variable X has probability distribution,

x	-1	0	1	3	7
$P(X=x)$	p	r	p	0.3	r

where p and r are probabilities.

Given that $E(X) = 1.95$

find the exact value of $E\left(\sqrt{X+1}\right)$ giving your answer in the form $a+b\sqrt{2}$ where a and b are rational.

(6)

Formula for E(X):

$E(x) = \Sigma x(P(X=x))$

In this case:

$E(x) = (-1 \times p) + (0 \times r) + (1 \times p) + 3(0.3) + 7(r) = 1.95$ given $E(X) = 1.95$

$-p + 0 + p + 0.9 + 7r = 1.95$

$7r = 1.05$

$r = 0.15$ M1A1

☆ Σ probabilities = 1 ∴

$2p + 2r + 0.3 = 1$

$2p + 0.6 = 1$ sub. $r = 0.15$

$2p = 0.4$

$p = 0.2$ M1A1

Let $Y = \sqrt{x+1}$ define a new variable.

Redraw the table including Y. The probabilities <u>don't</u> change!

x	-1	0	1	3	7
Y	0	1	$\sqrt{2}$	$\sqrt{4} = 2$	$\sqrt{8} = 2\sqrt{2}$
P(X=x)	0.2	0.15	0.2	0.3	0.15

to get Y's you just substitute each x into $\sqrt{x+1}$

Let's get E(Y): $E(Y) = 0(0.2) + 1(0.15) + \sqrt{2}(0.2) + 2(0.3) + 2\sqrt{2}(0.15)$ M1

$= 0.75 + 0.5\sqrt{2}$

as $Y = \sqrt{x+1}$, $E(Y) = E(\sqrt{x+1})$.

$\therefore E(\sqrt{x+1}) = 0.75 + 0.5\sqrt{2}$ A1

Question 3 continued

(Total for Question 3 is 6 marks)

4. Robin shoots 8 arrows at a target each day for 100 days.

The number of times he hits the target each day is summarised in the table below.

Number of hits	0	1	2	3	4	5	6	7	8
Frequency	1	10	30	34	17	4	2	0	2

Misha believes that these data can be modelled by a binomial distribution.

(a) State, in context, two assumptions that are implied by the use of this model. **(2)**

(b) Find an estimate for the proportion of arrows Robin shoots that hit the target. **(2)**

Misha calculates expected frequencies, to 2 decimal places, as follows.

Number of hits	0	1	2	3	4	5	6	7	8
Expected frequency	2.81	12.67	r	28.05	19.73	s	2.50	0.40	0.03

(c) Find the value of r and the value of s **(3)**

Misha correctly used a suitable test to assess her belief.

(d) (i) Explain why she used a test with 3 degrees of freedom. **(2)**

(ii) Complete the test using a 5% level of significance.
You should clearly state your hypotheses, test statistic, critical value and conclusion. **(6)**

(a) The probability of an arrow hitting the target is constant
The arrows are shot independantly B1 B1

★Remember Conditions for Binomial

(b) We need to use the data given to estimate probability.
use the Formula:

$$\frac{\Sigma(\text{\# of hits})\times(\text{frequency})}{\text{total tries}}$$

Substitute:

$$p = \frac{0\times1+1\times10+2\times30+3\times34+4\times17+5\times4+6\times2+7\times0+8\times2}{8\times100}$$ M1

$$= \frac{2.88}{8}$$

$p = 0.36$ A1

P 7 5 6 7 3 A 0 1 2 1 6

Question 4 continued

(c) To calculate the expected frequencies we will use binomial distribution.

$X \sim B(8, 0.36)$

$E_i = P(X=x) \times \Sigma$frequencies

In this case:

for $E(2)$: $P(X=2) = 0.24936$, total = 100 days M1 B1

$100 \times P(X=2) = 0.24936 \times 100 = 24.94 = r$ to 2dp

for $E(5)$: $P(X=5) = 0.08876$, total = 100 days

$100 \times P(X=5) = 8.88 = s$ to 2dp A1

(d) i. We need to pool so that all $E_i > 5$, ∴ we get 5 columns (pool col. 5-8 and 0-1) B1

We have 2 constraints as we estimated $p = 0.36$ in part (b)

∴ DoF = 5 - 2 = 3 ← DoF B1

columns, constraints

ii. Hypotheses

H_0: Binomial is a suitable model for the data

H_1: Binomial is not a suitable model for the data B1

Redraw the table:

Hits	0 or 1	2	3	4	5 to 8
O_i	11	30	34	17	8
E_i	15.48	24.94	28.05	19.73	11.80
$\frac{O_i^2}{E_i}$	7.82	36.09	41.21	14.65	5.42

for pooling you just add the columns you are pooling

M1

M1

Formula for χ^2:

$$\chi^2 = \Sigma \frac{O_i^2}{E_i} - N$$

Substitute:

$$\chi^2 = (7.82 + 36.09 + 41.21 + 14.65 + 5.42) - 100$$

(100 ← N)

$= 5.19$ to 3sf. test statistic. A1

Get c.v. from tables:

$\chi^2_3(0.05) = 7.815$ B1

3 DoF, 5% sign. level

Compare c.v. and test statistic:

$5.19 < 7.815$

Insufficient evidence to reject H_0.

Misha's Belief is supported, binomial is suitable for the data. A1

Question 4 continued

Question 4 continued

Question 4 continued

(Total for Question 4 is 15 marks)

TOTAL FOR FURTHER STATISTICS 1 IS 40 MARKS