AI SL 20.09 [113 marks]

1. [Maximum mark: 16] EXN.2.SL.TZ0.4

A dice manufacturer claims that for a novelty die he produces the probability of scoring the numbers 1 to 5 are all equal, and the probability of a 6 is two times the probability of scoring any of the other numbers.

(a) Find the probability of scoring a six when rolling the novelty die. [3]

Markscheme

* This sample question was produced by experienced DP mathematics senior examiners to aid teachers in preparing for external assessment in the new MAA course. There may be minor differences in formatting compared to formal exam papers.

Let the probability of scoring $1, \ \ldots, \ 5$ be p ,

$5p + 2p = 1 \Rightarrow p = \frac{1}{7}$ **(M1)(A1)**

Probability of $6 = \frac{2}{7}$ **A1**

[3 marks]

(b) Find the probability of scoring more than 2 sixes when this die is rolled 5 times. [4]

Markscheme

Let the number of sixes be X

$X \sim \mathrm{B}\left(5, \ \frac{2}{7}\right)$ **(M1)**

$\mathrm{P}(X > 2) = \mathrm{P}(X \geq 3)$ or $\mathrm{P}(X > 2) = 1 - \mathrm{P}(X \leq 2)$ **(M1)**

$= 0.145 \ (0.144701\ldots)$ **(M1)A1**

[4 marks]

To test the manufacture's claim one of the novelty dice is rolled 350 times and the numbers scored on the die are shown in the table below.

Number scored	Frequency
1	32
2	57
3	47
4	58
5	54
6	102

(c.i) Find the expected frequency for each of the numbers if the manufacturer's claim is true. [2]

Markscheme

Expected frequency is $350 \times p$ or $350 \times 2p$ **(M1)**

Number scored	Frequency	Expected frequency
1	32	50
2	57	50
3	47	50
4	58	50
5	54	50
6	102	100

A1

[2 marks]

A χ^2 goodness of fit test is to be used with a 5% significance level.

(c.ii) Write down the null and alternative hypotheses. [2]

Markscheme

H_0 : The manufacture’s claim is correct **A1**
H_1 : The manufacturer’s claim is not correct **A1**

[2 marks]

(c.iii) State the degrees of freedom for the test. [1]

Markscheme

Degrees of freedom $= 5$ **A1**

[1 mark]

(c.iv) Determine the conclusion of the test, clearly justifying your answer. [4]

Markscheme

p-value $= 0.0984$ $(0.0984037\ldots)$ **(M1)A1**

$0.0984 > 0.05$ **R1**

Hence insufficient evidence to reject the manufacture’s claim. **A1**

[4 marks]

2. [Maximum mark: 9] EXM.1.SL.TZ0.9

Six coins are tossed simultaneously 320 times, with the following results.

0 tail	5 times
1 tail	40 times
2 tails	86 times
3 tails	89 times
4 tails	67 times
5 tails	29 times
6 tails	4 times

At the 5% level of significance, test the hypothesis that all the coins are fair. [9]

Markscheme

Let H_0 be the hypothesis that all coins are fair, ***(C1)***

and let H_1 be the hypothesis that not all coins are fair. ***(C1)***

Let T be the number of tails obtained, T is binomially distributed. ***(M1)***

T	0	1	2	3	4	5	6
f_0	5	40	86	89	67	29	4
f_e	**5**	**30**	**75**	**100**	**75**	**30**	**5**

(A3)

Notes: Award ***(A2)*** if one entry on the third row is incorrect. Award ***(A1)*** if two entries on the third row are incorrect. Award ***(A0)*** if three or more entries on the third row are incorrect.

$$\chi^2_{\text{calc}} = \frac{(5-5)^2}{5} + \frac{(40-30)^2}{30} + \frac{(86-75)^2}{75} + \frac{(89-100)^2}{100} + \frac{(67-75)^2}{75} + \frac{(29-30)^2}{30} + \frac{(4-5)^2}{5}$$

$= 7.24$ ***(A1)***

Also $\chi^2_{0.05,\ 6} = 12.592$ ***(A1)***

Since 7.24 < 12.592, H_0 cannot be rejected. ***(R1)***

[9 marks]

3. [Maximum mark: 6] EXM.1.SL.TZ0.2

Kayla wants to measure the extent to which two judges in a gymnastics competition are in agreement. Each judge has ranked the seven competitors, as shown in the table, where 1 is the highest ranking and 7 is the lowest.

Competitor	A	B	C	D	E	F	G
Judge 1	1	2	3	3	5	6	6
Judge 2	2	3	1	4	5	5	7

(a) Calculate Spearman's rank correlation coefficient for this data. [5]

Markscheme

average equal ranks ***M1***

Competitor	A	B	C	D	E	F	G
Judge 1	1	2	3.5	3.5	5	6.5	6.5
Judge 2	2	3	1	4	5.5	5.5	7

A1A1

$r_s = 0.817$ ***A2***

[5 marks]

(b) State what conclusion Kayla can make from the answer in part (a). [1]

Markscheme

There is strong agreement between the two judges. ***R1***

[1 mark]

4. [Maximum mark: 5] 21M.1.SL.TZ1.6

Arriane has geese on her farm. She claims the mean weight of eggs from her black geese is less than the mean weight of eggs from her white geese.

She recorded the weights of eggs, in grams, from a random selection of geese. The data is shown in the table.

Weights of eggs from black geese	136	134	142	141	128	126
Weights of eggs from white geese	135	138	141	140	136	134

In order to test her claim, Arriane performs a t-test at a 10% level of significance. It is assumed that the weights of eggs are normally distributed and the samples have equal variances.

(a) State, in words, the null hypothesis. [1]

Markscheme

EITHER

H_0: The population mean weight of eggs from (her/the) black geese is equal to/the same as the population mean weight of eggs from (her/the) white geese.

OR

H_0: The population mean weight of eggs from (her/the) black geese is not less than the population mean weight of eggs from (her/the) white geese.

A1

Note: Reference to the "population mean weight" must be explicit for the ***A1*** to be awarded. The term “population” can be implied by use of “all” or “on average” or “generally” when relating to the weight of eggs e.g. “the mean weight of eggs for all (her/the) black geese”.
Award ***A0*** if reference is made to the mean weights from the sample or the

table.
Award ***A0*** for a null hypothesis written in symbolic form.

[1 mark]

(b) Calculate the p-value for this test. [2]

Markscheme

p-value $= 0.177\ \ (0.176953\ldots)$ ***A2***

Note: Award ***A1*** for an answer of $0.18221\ldots$, from "unpooled" settings on GDC.

[2 marks]

(c) State whether the result of the test supports Arriane's claim.
Justify your reasoning. [2]

Markscheme

$0.177 > 0.1$ ***R1***

(insufficient evidence to reject H_0)

Arriane's claim is not supported by the evidence ***A1***

Note: Accept $p > 0.1$ or $p >$ *significance level* provided p is explicitly seen in part (b). Award ***A1*** only if reference is specifically made to Arriane's claim. Do not award ***R0A1***.

[2 marks]

5. [Maximum mark: 6] 21M.1.SL.TZ2.11

A newspaper vendor in Singapore is trying to predict how many copies of *The Straits Times* they will sell. The vendor forms a model to predict the number of copies sold each weekday. According to this model, they expect the same number of copies will be sold each day.

To test the model, they record the number of copies sold each weekday during a particular week. This data is shown in the table.

Day	Monday	Tuesday	Wednesday	Thursday	Friday
Number of copies sold	74	97	91	86	112

A goodness of fit test at the 5% significance level is used on this data to determine whether the vendor's model is suitable.

The critical value for the test is 9.49 and the hypotheses are

H_0 : The data satisfies the model.
H_1 : The data does not satisfy the model.

(a) Find an estimate for how many copies the vendor expects to sell each day. [1]

Markscheme

$\left(\frac{74+97+91+86+112}{5}\right) = 92$ ***A1***

[1 mark]

(b.i) Write down the degrees of freedom for this test. [1]

Markscheme

4 ***A1***

[1 mark]

(b.ii) Write down the conclusion to the test. Give a reason for your answer. [4]

Markscheme

$\chi^2{}_{\text{calc}} = 8.54\ \ (8.54347\ldots)$ **OR** p-value$= 0.0736\ \ (0.0735802\ldots)$ ***A2***

$8.54 < 9.49$ **OR** $0.0736 > 0.05$ ***R1***

therefore there is insufficient evidence to reject H_0 ***A1***

(i.e. the data satisfies the model)

Note: Do not award ***R0A1***. Accept "accept" or "do not reject" in place of "insufficient evidence to reject".

Award the ***R1*** for comparing their p-value with 0.05 or their χ^2 value with 9.49 and then ***FT*** their final conclusion.

[4 marks]

6. [Maximum mark: 6] 21M.1.SL.TZ2.8

At Springfield University, the weights, in kg, of 10 chinchilla rabbits and 10 sable rabbits were recorded. The aim was to find out whether chinchilla rabbits are generally heavier than sable rabbits. The results obtained are summarized in the following table.

Weight of chinchilla rabbits, kg	4.9	4.2	4.1	4.4	4.3	4.6	4.0	4.7	4.5	4.4
Weight of sable rabbits, kg	4.2	4.1	4.1	4.2	4.5	4.4	4.5	3.9	4.2	4.0

A t-test is to be performed at the 5% significance level.

(a) Write down the null and alternative hypotheses. [2]

Markscheme

(let $\mu_c =$ population mean for chinchilla rabbits, $\mu_s =$ population mean for sable rabbits)

$\mathrm{H}_0 : \mu_c = \mu_s$ ***A1***

$\mathrm{H}_1 : \mu_c > \mu_s$ ***A1***

Note: Accept an equivalent statement in words, must include mean and reference to "**population** mean" / "mean for **all** chinchilla rabbits" for the first ***A1*** to be awarded.

Do not accept an imprecise *"the means are equal"*.

[2 marks]

(b) Find the p-value for this test. [2]

Markscheme

p-value $= 0.0408 \quad (0.0408065\ldots)$ ***A2***

Note: Award ***A1*** for an answer of $0.041565\ldots$, from “unpooled” settings on GDC.

[2 marks]

(c) Write down the conclusion to the test. Give a reason for your answer. [2]

Markscheme

$0.0408 < 0.05$. ***R1***

(there is sufficient evidence to) reject (or not accept) H_0 ***A1***

(there is sufficient evidence to suggest that chinchilla rabbits are heavier than sable rabbits)

Note: Do not award ***R0A1***. Accept ‘accept H_1’.

[2 marks]

7. [Maximum mark: 18] 21M.2.SL.TZ1.1

As part of his mathematics exploration about classic books, Jason investigated the time taken by students in his school to read the book *The Old Man and the Sea*. He collected his data by stopping and asking students in the school corridor, until he reached his target of 10 students from **each** of the literature classes in his school.

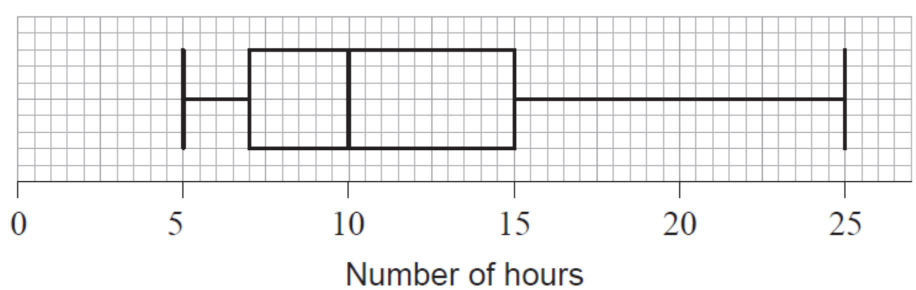
(a) State which of the two sampling methods, systematic or quota, Jason has used. [1]

Markscheme

Quota sampling ***A1***

[1 mark]

Jason constructed the following box and whisker diagram to show the number of hours students in the sample took to read this book.



(b) Write down the median time to read the book. [1]

Markscheme

10 (hours) ***A1***

[1 mark]

(c) Calculate the interquartile range. [2]

Markscheme

$15 - 7$ ***(M1)***

Note: Award ***M1*** for 15 and 7 seen.

8 ***A1***

[2 marks]

Mackenzie, a member of the sample, took 25 hours to read the novel. Jason believes Mackenzie's time is not an outlier.

(d) Determine whether Jason is correct. Support your reasoning. [4]

Markscheme

indication of a valid attempt to find the upper fence ***(M1)***

$15 + 1.5 \times 8$

27 ***A1***

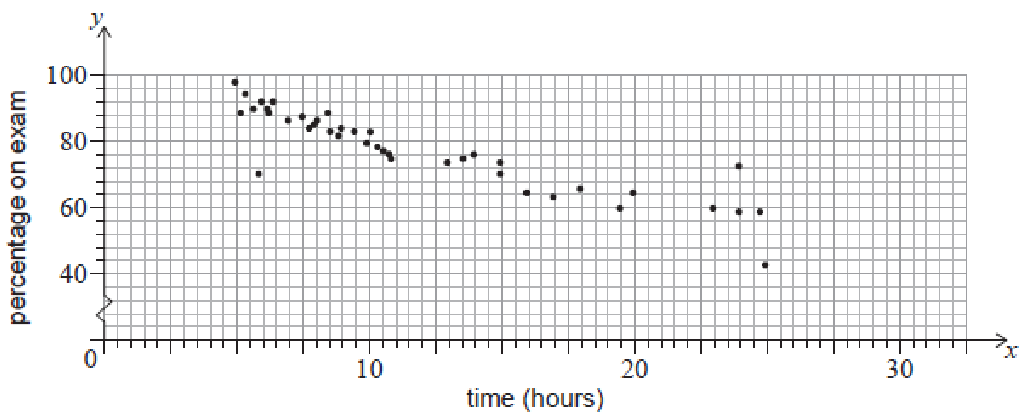
$25 < 27$ (accept equivalent answer in words) ***R1***

Jason is correct ***A1***

Note: Do not award ***R0A1***. Follow through **within** this part from *their* 27, but only if their value is supported by a valid attempt **or** clearly and correctly explains what their value represents.

[4 marks]

For each student interviewed, Jason recorded the time taken to read *The Old Man and the Sea* (x), measured in hours, and paired this with their percentage score on the final exam (y). These data are represented on the scatter diagram.



(e) Describe the correlation. [1]

Markscheme

"negative" seen ***A1***

Note: Strength cannot be inferred visually; ignore "strong" or "weak".

[1 mark]

Jason correctly calculates the equation of the regression line y on x for these students to be

$y = -1.54x + 98.8.$

He uses the equation to estimate the percentage score on the final exam for a student who read the book in 1.5 hours.

(f) Find the percentage score calculated by Jason. [2]

Markscheme

correct substitution *(M1)*

$y = -1.54 \times 1.5 + 98.8$

$96.5(\%)$ (96.49) *A1*

[2 marks]

(g) State whether it is valid to use the regression line y on x for Jason's estimate. Give a reason for your answer. [2]

Markscheme

not reliable *A1*

extrapolation **OR** outside the given range of the data *R1*

Note: Do not award ***A1R0***. Only accept reasoning that includes reference to the range of the data. Do not accept a contextual reason such as 1.5 hours is too short to read the book.

[2 marks]

Jason found a website that rated the 'top 50' classic books. He randomly chose eight of these classic books and recorded the number of pages. For example, Book H is rated $44\mathrm{th}$ and has 281 pages. These data are shown in the table.

Book	A	B	C	D	E	F	G	H
Number of pages (n)	4215	863	585	1225	366	209	624	281
Top 50 rating (t)	1	2	5	7	13	22	40	44

Jason intends to analyse the data using Spearman's rank correlation coefficient, r_s.

(h) Copy and complete the information in the following table. [2]

Book	A	B	C	D	E	F	G	H
Rank – Number of pages	1							
Rank – Top 50 Rating	1							

Markscheme

	Book							
	A	B	C	D	E	F	G	H
Rank – Number of pages	1	3	5	2	6	8	4	7
Rank – Top 50 Rating	1	2	3	4	5	6	7	8

A1A1

Note: Do not award ***A1*** for correct ranks for 'number of pages'. Award ***A1*** for correct ranks for 'top 50 rating'.

[2 marks]

(i.i) Calculate the value of r_s. [2]

Markscheme

$0.714\ \ (0.714285\ldots)$ ***A2***

Note: ***FT*** from their table.

[2 marks]

(i.ii) Interpret your result. [1]

Markscheme

EITHER

there is a (strong/moderate) positive association between the number of pages and the top 50 rating. ***A1***

OR

there is a (strong/moderate) agreement between the rank order of number of pages and the rank order top 50 rating. ***A1***

OR

there is a (strong/moderate) positive (linear) correlation between the rank order of number of pages and the rank order top 50 rating. ***A1***

Note: Follow through from their value of r_s.

[1 mark]

8. [Maximum mark: 13] 21M.2.SL.TZ1.4

The stopping distances for bicycles travelling at $20\ \mathrm{km\,h^{-1}}$ are assumed to follow a normal distribution with mean $6.76\ \mathrm{m}$ and standard deviation $0.12\ \mathrm{m}$.

Under this assumption, find, correct to four decimal places, the probability that a bicycle chosen at random travelling at $20\ \mathrm{km\,h^{-1}}$ manages to stop

(a.i) in less than $6.5\ \mathrm{m}$. [2]

Markscheme

evidence of correct probability ***(M1)***

e.g sketch **OR** correct probability statement, $\mathrm{P}(X < 6.5)$

0.0151 ***A1***

[2 marks]

(a.ii) in more than $7\ \mathrm{m}$. [1]

Markscheme

0.0228 ***A1***

Note: Answers should be given to 4 decimal place.

[1 mark]

1000 randomly selected bicycles are tested and their stopping distances when travelling at $20\ \mathrm{km\,h^{-1}}$ are measured.

Find, correct to four significant figures, the expected number of bicycles tested that stop between

(b.i) $6.5\ \text{m}$ and $6.75\ \text{m}$. [2]

Markscheme

multiplying **their** probability by 1000 ***(M1)***

451.7 ***A1***

[2 marks]

(b.ii) $6.75\ \text{m}$ and $7\ \text{m}$. [1]

Markscheme

510.5 ***A1***

Note: Answers should be given to 4 sf.

[1 mark]

The measured stopping distances of the 1000 bicycles are given in the table.

Measured stopping distance	**Number of bicycles**
Less than 6.5m	12
Between 6.5m and 6.75m	428
Between 6.75m and 7m	527
More than 7m	33

It is decided to perform a χ^2 goodness of fit test at the 5% level of significance to decide whether the stopping distances of bicycles travelling at $20\ \mathrm{km\,h^{-1}}$ can be modelled by a normal distribution with mean $6.76\ \mathrm{m}$ and standard deviation $0.12\ \mathrm{m}$.

(c) State the null and alternative hypotheses. [2]

Markscheme

$\mathrm{H_0}$: stopping distances can be modelled by $\mathrm{N}\left(6.76,\ 0.12^2\right)$
$\mathrm{H_1}$: stopping distances cannot be modelled by $\mathrm{N}\left(6.76,\ 0.12^2\right)$ ***A1A1***

Note: Award ***A1*** for correct $\mathrm{H_0}$, including reference to the mean and standard deviation. Award ***A1*** for the negation of their $\mathrm{H_0}$.

[2 marks]

(d) Find the p-value for the test. [3]

Markscheme

15.1 or 22.8 seen ***(M1)***

$0.0727\ \ (0.0726542\ldots,\ 7.27\%)$ ***A2***

[3 marks]

(e) State the conclusion of the test. Give a reason for your answer. [2]

Markscheme

$0.05 < 0.0727$ ***R1***

there is insufficient evidence to reject $\mathrm{H_0}$ (or "accept $\mathrm{H_0}$") ***A1***

Note: Do not award ***R0A1***.

[2 marks]

9. [Maximum mark: 6] 20N.1.SL.TZ0.T_10

On 90 journeys to his office, Isaac noted whether or not it rained. He also recorded his journey time to the office, and classified each journey as short, medium or long.

Of the 90 journeys to the office, there were 3 short journeys when it rained, 22 medium journeys when it rained, and 15 long journeys when it rained. There were also 14 short journeys when it did not rain.

Isaac carried out a χ^2 test at the 5% level of significance on these data, looking at the weather and the types of journeys.

(a) Write down H_0, the null hypothesis for this test. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

type of journey and whether it rained are independent ***(A1) (C1)***

Note: Accept "there is no association" or "not dependent". Do not accept "not related" or "not correlated". Accept equivalent terms for 'type of journey'.

[1 mark]

(b) Find the expected number of short trips when it rained. [3]

Markscheme

$\frac{17}{90} \times \frac{40}{90} \times 90$ **OR** $\frac{17 \times 40}{90}$ ***(A1)(M1)***

Note: Award ***(A1)*** for 17 or 40 seen. Award ***(M1)*** for $\frac{17}{90} \times \frac{40}{90} \times 90$ **OR**

$\frac{17\times 40}{90}$ seen.

7.56 $\left(7.55555\ldots,\ \frac{68}{9}\right)$ ***(A1) (C3)***

[3 marks]

(c) The p-value for this test is 0.0206.

State the conclusion to Isaac's test. Justify your reasoning. [2]

Markscheme

reject (do not accept) H_0 ***(A1)***

OR

type of journey and whether it rained are not independent ***(A1)***

Note: Follow through from part (a) for their phrasing of the null hypothesis.

$0.0206 < 0.05$ ***(R1) (C2)***

Note: A comparison must be seen, either numerically or in words (e.g. p-value < significance level). Do not award ***(R0)(A1).***

[2 marks]

10. [Maximum mark: 15] 19N.2.SL.TZ0.T_1

Casanova restaurant offers a set menu where a customer chooses **one** of the following meals: pasta, fish or shrimp.

The manager surveyed 150 customers and recorded the customer's age and chosen meal. The data is shown in the following table.

	Pasta	Fish	Shrimp	Total
Adults	24	25	32	81
Children	20	14	35	69
Total	44	39	67	150

A χ^2 test was performed at the 10% significance level. The critical value for this test is 4.605.

(a) State H_0, the null hypothesis for this test. [1]

Markscheme

$(\mathrm{H}_0:)$ choice of meal is independent of age (or equivalent) ***(A1)***

Note: Accept "not associated" or "not dependent" instead of independent. In lieu of "age", accept an equivalent alternative such as "being a child or adult".

[1 mark]

(b) Write down the number of degrees of freedom. [1]

Markscheme

2 ***(A1)***

[1 mark]

(c) Show that the expected number of children who chose shrimp is 31, correct to two significant figures. [2]

Markscheme

$\frac{69}{150} \times \frac{67}{150} \times 150$ **OR** $\frac{69 \times 67}{150}$ ***(M1)***

Note: Award ***(M1)*** for correct substitution into expected frequency formula.

$30.82\ \ (30.8)$ ***(A1)***

31 ***(AG)***

Note: Both an unrounded answer that rounds to the given answer and rounded answer must be seen for the ***(A1)*** to be awarded.

[2 marks]

Write down

(d.i) the χ_2 statistic. [2]

Markscheme

$(\chi^2_{\text{calc}} =)\ 2.66\ \ (2.657537\ldots)$ ***(G2)***

[2 marks]

(d.ii) the p-value. [1]

Markscheme

(p-value $=$) $0.265\ (0.264803\ldots)$ ***(G1)***

Note: Award ***(G0)(G2)*** if the χ^2 statistic is missing or incorrect and the p-value is correct.

[1 mark]

(e) State the conclusion for this test. Give a reason for your answer. [2]

Markscheme

$0.265 > 0.10$ ***OR*** $2.66 < 4.605$ ***(R1)*****(ft)**

the null hypothesis is not rejected ***(A1)*****(ft)**

OR

the choice of meal is independent of age (or equivalent) ***(A1)*****(ft)**

Note: Award ***(R1)*****(ft)**) for a correct comparison of either their χ^2 statistic to the χ^2 critical value or their p-value to the significance level.
Condone "accept" in place of "not reject".
Follow through from parts (a) and (d).

Do not award ***(A1)*****(ft)*****(R0)***.

[2 marks]

A customer is selected at random.

(f.i) Calculate the probability that the customer is an adult. [2]

Markscheme

$\frac{81}{150}\left(\frac{27}{50},\ 0.54,\ 54\%\right)$ ***(A1)(A1)(G2)***

Note: Award ***(A1)*** for numerator, ***(A1)*** for denominator.

[2 marks]

(f.ii) Calculate the probability that the customer is an adult or that the customer chose shrimp. [2]

Markscheme

$\frac{116}{150}$ $\left(\frac{58}{75},\ 0.773,\ 0.773333\ldots,\ 77.3\%\right)$ ***(A1)(A1)(G2)***

Note: Award ***(A1)*** for numerator, ***(A1)*** for denominator.

[2 marks]

(f.iii) Given that the customer is a child, calculate the probability that they chose pasta or fish. [2]

Markscheme

$\frac{34}{69}$ $(0.493,\ 0.492753\ldots,\ 49.3\%)$ ***(A1)(A1)(G2)***

Note: Award ***(A1)*** for numerator, ***(A1)*** for denominator.

[2 marks]

11. [Maximum mark: 13] 19M.2.SL.TZ1.T_3

A survey was conducted on a group of people. The first question asked how many pets they each own. The results are summarized in the following table.

Number of pets owned	0	1	2	3	4	5
Number of people	20	45	40	30	20	5

(a) Write down the total number of people, from this group, who are **pet owners**. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

140 ***(A1)***

[1 mark]

(b) Write down the modal number of pets. [1]

Markscheme

1 ***(A1)***

[1 mark]

(c.i) For these data, write down the median number of pets. [1]

Markscheme

2 ***(A1)***

[1 mark]

(c.ii) For these data, write down the lower quartile. [1]

Markscheme

1 ***(A1)***

[1 mark]

(c.iii) For these data, write down the upper quartile. [1]

Markscheme

3 ***(A1)***

[1 mark]

The second question asked each member of the group to state their age and preferred pet. The data obtained is organized in the following table.

	Age	
Preferred pet	**Teenager**	**Non-teenager**
cat	23	32
dog	35	23
bird	16	13
other	11	7

(d) Write down the ratio of teenagers to non-teenagers in its simplest form. [1]

Markscheme

17:15 **OR** $\frac{17}{15}$ ***(A1)***

Note: Award ***(A0)*** for 85:75 or 1.13:1.

[1 mark]

A χ^2 test is carried out at the 10 % significance level.

(e.i) State the null hypothesis. [1]

Markscheme

preferred pet is independent of "whether or not the respondent was a teenager" or "age category" ***(A1)***

Note: Accept there is no association between pet and age. Do not accept "not related" or "not correlated" or "influenced".

[1 mark]

(e.ii) State the alternative hypothesis. [1]

Markscheme

preferred pet is not independent of age ***(A1)*****(ft)**

Note: Follow through from part (e)(i) *i.e.* award ***(A1)*****(ft)** if their alternative hypothesis is the negation of their null hypothesis. Accept "associated" or "dependent".

[1 mark]

(f) Write down the number of degrees of freedom for this test. [1]

Markscheme

3 ***(A1)***

[1 mark]

(g) Calculate the expected number of teenagers that prefer cats. [2]

Markscheme

$\frac{85 \times 55}{160}$ **OR** $\frac{85}{160} \times \frac{55}{160} \times 160$ ***(M1)***

29.2 (29.2187…) ***(A1)(G2)***

[2 marks]

(i) State the conclusion for this test. Give a reason for your answer. [2]

Markscheme

0.208 > 0.1 ***(R1)***

accept null hypothesis **OR** fail to reject null hypothesis ***(A1)*(ft)**

Note: Award ***(R1)*** for a correct comparison of their p-value to the significance level, award ***(A1)*(ft)** for the correct result from that comparison. Accept "p-value > 0.1" as part of the comparison but only if their p-value is explicitly seen in part (h). Follow through from their answer to part (h). Do not award ***(R0)(A1)***.

[2 marks]