

# chi squared [54 marks]

As part of a study into healthy lifestyles, Jing visited Surrey Hills University. Jing recorded a person's position in the university and how frequently they ate a salad. Results are shown in the table.

	Salad meals per week			
	0	1–2	3–4	>4
Students	45	26	18	6
Professors	15	8	5	12
Staff and Administration	16	13	10	6

Jing conducted a  $\chi^2$  test for independence at a 5 % level of significance.

1a. State the null hypothesis.

[1 mark]

## Markscheme

number of salad meals per week is independent of a person's position in the university **A1**

**Note:** Accept "not associated" instead of independent.

[1 mark]

1b. Calculate the  $p$ -value for this test.

[2 marks]

## Markscheme

0.0201 (0.0201118...) **A2**

[2 marks]

1c. State, giving a reason, whether the null hypothesis should be accepted. [2 marks]

# Markscheme

$0.0201 < 0.05$       **R1**

the null hypothesis is rejected      **A1**

**Note:** Award **(R1)** for a correct comparison of their  $p$ -value to the test level, award **(A1)** for the correct interpretation from that comparison.

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

**[2 marks]**

Manny and Annabelle, mathematics teachers at Burnham High School, give their students the same examination. A random sample of the examination scores were collected from each of their classes.

Examination scores from Manny's class	76	77	82	84	88	90	91	98
Examination scores from Annabelle's class	68	79	81	89	91	92	92	95

Annabelle uses these scores to conduct a two-tailed  $t$ -test to compare the means of the two classes, at the 5% level of significance. It is assumed the examination scores for both classes have the same variance and are normally distributed.

The null hypothesis is  $\mu_1 = \mu_2$ , where  $\mu_1$  is the mean examination score from Manny's class and  $\mu_2$  is the mean examination score from Annabelle's class.

2a. Write down the alternative hypothesis.

**[1 mark]**

# Markscheme

$(H_1 : ) \mu_1 \neq \mu_2$       **A1**

**Note:** Accept an equivalent statement in words referring to  $\mu_1$  and  $\mu_2$  as defined in the question.

**[1 mark]**

2b. Find the  $p$ -value for this test. Give your answer correct to five decimal places.      **[2 marks]**

# Markscheme

0.97652 (0.976516...)

**A2**

**[2 marks]**

Annabelle concludes there is insufficient evidence to reject the null hypothesis.

2c. State whether Annabelle's conclusion is correct. Give a reason for your answer. **[2 marks]**

# Markscheme

$0.97652 > 0.05$  ( $0.977 > 0.05$ )

**R1**

Annabelle's conclusion is correct.

**A1**

**Note:** Do not award **ROA1**. Answer must reference Annabelle's conclusion; do not accept an answer, without context, of "fail to reject  $H_0$ " for the **A1** mark.

**[2 marks]**

A group of 1280 students were asked which electronic device they preferred. The results per age group are given in the following table.

Preferred device	Age			Total
	11–13	14–16	17–18	
Laptop	143	160	153	456
Tablet	205	224	131	560
Mobile phone	72	128	64	264
Total	420	512	348	1280

A student from the group is chosen at random. Calculate the probability that the student

3a. prefers a tablet.

**[2 marks]**

## Markscheme

$$\frac{560}{1280} \left( \frac{7}{16}, 0.4375 \right) \quad \mathbf{A1A1}$$

**Note:** Award **A1** for correct numerator, **A1** for correct denominator.

**[2 marks]**

3b. is 11–13 years old and prefers a mobile phone.

**[2 marks]**

## Markscheme

$$\frac{72}{1280} \left( \frac{9}{160}, 0.05625 \right) \quad \mathbf{A1A1}$$

**Note:** Award **A1** for correct numerator, **A1** for correct denominator.

**[2 marks]**

3c. prefers a laptop **given that** they are 17–18 years old.

**[2 marks]**

## Markscheme

$$\frac{153}{348} \left( \frac{51}{116}, 0.439655\dots \right) \quad \mathbf{A1A1}$$

**Note:** Award **A1** for correct numerator, **A1** for correct denominator.

**[2 marks]**

3d. prefers a tablet or is 14–16 years old.

**[3 marks]**

## Markscheme

$$160 + 224 + 128 + 205 + 131 \text{ OR } 560 + 512 - 224 \quad (\mathbf{M1})$$

$$\frac{848}{1280} \left( \frac{53}{80}, 0.6625 \right) \quad \mathbf{A1A1}$$

**Note:** Award **A1** for correct denominator (1280) seen, **(M1)** for correct calculation of the numerator, **A1** for the correct answer.

**[3 marks]**

A  $\chi^2$  test for independence was performed on the collected data at the 1% significance level. The critical value for the test is 13.277.

3e. State the null and alternative hypotheses.

**[1 mark]**

## Markscheme

$H_0$  : the variables are independent

$H_1$  : the variables are dependent **A1**

**Note:** Award **A1** for for both hypotheses correct. Do not accept “not correlated” or “not related” in place of “independent”.

**[1 mark]**

3f. Write down the number of degrees of freedom.

**[1 mark]**

## Markscheme

4 **A1**

**[1 mark]**

3g. Write down the  $\chi^2$  test statistic.

**[2 marks]**

## Markscheme

$(\chi^2 =)$  23.3 (23.3258...) **A2**

**[2 marks]**

3h. Write down the  $p$ -value.

**[1 mark]**

## Markscheme

0.000109 (0.000108991...) **OR**  $1.09 \times 10^{-4}$  **A1**

**[1 mark]**

3i. State the conclusion for the test in context. Give a reason for your answer.

**[2 marks]**

## Markscheme

**EITHER**

$23.3 > 13.277$  **R1**

**OR**

$0.000109 < 0.01$  **R1**

**THEN**

(there is sufficient evidence to accept  $H_1$  that) preferred device and age group are not independent **A1**

**Note:** For the final **A1** the answer must be in context. Do not award **A1R0**.

**[2 marks]**

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  $\chi^2$  test was performed at the 10% significance level. The critical value for this test is 4.605.

4a. State  $H_0$ , the null hypothesis for this test.

[1 mark]

## Markscheme

( $H_0$ ): 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]

4b. Write down the number of degrees of freedom.

[1 mark]

## Markscheme

2 **(A1)**

[1 mark]

4c. Show that the expected number of children who chose shrimp is 31, correct to two significant figures.

[2 marks]

## Markscheme

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

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

$$30.82 \text{ (30.8)} \quad (A1)$$

$$31 \quad (AG)$$

**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

4d. the  $\chi^2$  statistic.

**[2 marks]**

## Markscheme

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

**[2 marks]**

4e. the  $p$ -value.

**[1 mark]**

## Markscheme

$$(p\text{-value} =) 0.265 \text{ (0.264803...)} \quad (G1)$$

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

**[1 mark]**

4f. State the conclusion for this test. Give a reason for your answer.

**[2 marks]**



## 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  $\chi^2$  statistic to the  $\chi^2$  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.

4g. Calculate the probability that the customer is an adult.

**[2 marks]**

## Markscheme

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

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

**[2 marks]**

4h. Calculate the probability that the customer is an adult or that the customer chose shrimp.

**[2 marks]**

## Markscheme

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

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

**[2 marks]**

4i. Given that the customer is a child, calculate the probability that they chose pasta or fish.

**[2 marks]**

# Markscheme

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

**Note:** Award (A1) for numerator, (A1) for denominator.  
[2 marks]

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

- 5a. Write down the total number of people, from this group, who are **pet owners**. [1 mark]

# Markscheme

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

140 (A1)  
[1 mark]

- 5b. Write down the modal number of pets. [1 mark]

# Markscheme

1 (A1)  
[1 mark]

- 5c. For these data, write down the median number of pets. [1 mark]

# Markscheme

2 (A1)  
[1 mark]

5d. For these data, write down the lower quartile.

[1 mark]

# Markscheme

1 (A1)  
[1 mark]

5e. For these data, write down the upper quartile.

[1 mark]

# 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.

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

5f. Write down the ratio of teenagers to non-teenagers in its simplest form. [1 mark]

# Markscheme

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

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

[1 mark]

A  $\chi^2$  test is carried out at the 10 % significance level.

5g. State the null hypothesis.

[1 mark]

# 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]

5h. State the alternative hypothesis.

[1 mark]

# 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]

5i. Write down the number of degrees of freedom for this test.

[1 mark]

# Markscheme

3 (A1)

[1 mark]

5j. Calculate the expected number of teenagers that prefer cats.

[2 marks]

## Markscheme

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

29.2 (29.2187...) (A1)(G2)

[2 marks]

5k. State the conclusion for this test. Give a reason for your answer.

[2 marks]

## 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]