

# Chi Sq

215 min  
203 marks

1. Some of the customers in each café were given survey forms to complete to find out if they were satisfied with the standard of service they received.

	Pete's Eats	Alan's Diner	Sarah's Snackbar	Total
Dissatisfied	16	8	16	40
Satisfied	26	20	34	80
Total	42	28	50	120

One of the survey forms was chosen at random, find the probability that

- (a) the form showed "Dissatisfied"; (2)
- (b) the form showed "Satisfied" and was completed at Sarah's Snackbar; (2)
- (c) the form showed "Dissatisfied", given that it was completed at Alan's Diner. (2)

A  $\chi^2$  test at the 5 % significance level was carried out to determine whether there was any difference in the level of customer satisfaction in each of the cafés.

- (d) Write down the null hypothesis,  $H_0$ , for the  $\chi^2$  test. (1)

(e) Write down the number of degrees of freedom for the test. (1)

(f) Using your graphic display calculator, find  $\chi^2_{calc}$ . (2)

(g) State, giving a reason, the conclusion to the test. (2)  
(Total 12 marks)

2. A university required all Science students to study one language for one year. A survey was carried out at the university amongst the 150 Science students. These students all studied one of either French, Spanish or Russian. The results of the survey are shown below.

	French	Spanish	Russian
Female	9	29	12
Male	31	40	29

Ludmila decides to use the  $\chi^2$  test at the 5 % level of significance to determine whether the choice of language is independent of gender.

(a) State Ludmila's null hypothesis. (1)

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

(c) Find the expected frequency for the females studying Spanish. (2)

(d) Use your graphic display calculator to find the  $\chi^2$  test statistic for this data. (2)

(e) State whether Ludmila accepts the null hypothesis. Give a reason for your answer.

(2)

(Total 8 marks)

3. The same 100 students are also asked how many meals on average they have per day. The data collected is organized in the following table.

	3 or fewer meals per day	4 or 5 meals per day	More than 5 meals per day	Total
Male	15	25	15	55
Female	12	20	13	45
Total	27	45	28	100

A  $\chi^2$  test is carried out at the 5 % level of significance.

(a) Write down the null hypothesis,  $H_0$ , for this test.

(1)

(b) Write down the number of degrees of freedom for this test.

(1)

(c) Write down the critical value for this test.

(1)

(d) Show that the expected number of females that have more than 5 meals per day is 13, correct to the nearest integer.

(2)

(e) Use your graphic display calculator to find the  $\chi^2_{\text{calc}}$  for this data.

(2)

(f) Decide whether  $H_0$  must be accepted. Justify your answer.

(2)

(Total 9 marks)

4. A manufacturer claims that fertilizer has an effect on the height of rice plants. He measures the height of fertilized and unfertilized plants. The results are given in the following table.

Plant height	Fertilized plants	Unfertilized plants
> 75 cm	115	80
50 – 75 cm	45	65
< 50 cm	20	35

A chi-squared test is performed to decide if the manufacturer's claim is justified at the **1 %** level of significance.

- (a) Write down the null and alternate hypotheses for this test. (2)
- (b) For the number of fertilized plants with height greater than 75 cm, show that the expected value is 97.5. (3)
- (c) Write down the value of  $\chi^2_{calc}$ . (2)
- (d) Write down the number of degrees of freedom. (1)
- (e) Write down the critical value of  $\chi^2$ , at the **1 %** level of significance. (1)
- (f) Is the manufacturer's claim justified? Give a reason for your answer. (2)

(Total 11 marks)

5. In a study on 100 students there seemed to be a difference between males and females in their choice of favourite car colour. The results are given in the table below.  
A  $\chi^2$  test was conducted.

	Blue	Red	Green
Males	14	6	8
Females	31	24	17

- (a) Write down the total number of male students. (1)

- (b) Show that the expected frequency for males, whose favourite car colour is blue, is 12.6. (2)

The calculated value of  $\chi^2$  is 1.367.

- (c) (i) Write down the null hypothesis for this test.  
(ii) Write down the number of degrees of freedom.  
(iii) Write down the critical value of  $\chi^2$  at the 5 % significance level.  
(iv) Determine whether the null hypothesis should be accepted. Give a reason for your answer.

(5)  
(Total 8 marks)

6. A survey of 400 people is carried out by a market research organization in two different cities, Buenos Aires and Montevideo. The people are asked which brand of cereal they prefer out of Chocos, Zucos or Fruti. The table below summarizes their responses.

	Chocos	Zucos	Fruti	Total
Buenos Aires	43	85	62	190
Montevideo	57	35	118	210
Total	100	120	180	400

- (a) One person is chosen at random from those surveyed. Find the probability that this person
- (i) does not prefer Zucos;
  - (ii) prefers Chocos, given that they live in Montevideo.
- (4)
- (b) Two people are chosen at random from those surveyed. Find the probability that they both prefer Fruti.
- (3)

The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed.

- (c) State the null hypothesis.
- (1)
- (d) State the number of degrees of freedom.
- (1)
- (e) Show that the expected frequency for the number of people who live in Montevideo and prefer Zucos is 63.
- (2)
- (f) Write down the chi-squared statistic for this data.
- (2)
- (g) State whether the market research organization would accept the null hypothesis. Clearly justify your answer.
- (2)

**(Total 15 marks)**

7. Jorge conducted a survey of 200 drivers. He asked two questions:

How long have you had your driving licence?  
Do you wear a seat belt when driving?

The replies are summarized in the table below.

	<b>Wear a seat belt</b>	<b>Do not wear a seat belt</b>
<b>Licence less than 2 years</b>	38	42
<b>Licence between 2 and 15 years</b>	30	45
<b>Licence more than 15 years</b>	30	15

- (a) Jorge applies a  $\chi^2$  test at the 5 % level to investigate whether wearing a seat belt is associated with the time a driver has had their licence.
- (i) Write down the null hypothesis,  $H_0$ .
  - (ii) Write down the number of degrees of freedom.
  - (iii) Show that the expected number of drivers that wear a seat belt and have had their driving licence for more than 15 years is 22, correct to the nearest whole number.
  - (iv) Write down the  $\chi^2$  test statistic for this data.
  - (v) Does Jorge accept  $H_0$ ? Give a reason for your answer.

(8)

- (b) Consider the 200 drivers surveyed. One driver is chosen at random. Calculate the probability that
- (i) this driver wears a seat belt;
  - (ii) the driver does not wear a seat belt, **given that** the driver has held a licence for more than 15 years.

(4)

- (c) Two drivers are chosen at random. Calculate the probability that
- (i) both wear a seat belt.
  - (ii) at least one wears a seat belt.

(6)  
(Total 18 marks)

8. Manuel conducts a survey on a random sample of 751 people to see which television programme type they watch most from the following: Drama, Comedy, Film, News. The results are as follows.

	Drama	Comedy	Film	News
Males under 25	22	65	90	35
Males 25 and over	36	54	67	17
Females under 25	22	59	82	15
Females 25 and over	64	39	38	46

Manuel decides to ignore the ages and to test at the 5% level of significance whether the most watched programme type is independent of **gender**.

- (a) Draw a table with 2 rows and 4 columns of data so that Manuel can perform a chi-squared test. (3)
- (b) State Manuel's null hypothesis and alternative hypothesis. (1)
- (c) Find the expected frequency for the number of females who had "Comedy" as their most-watched programme type. Give your answer to the nearest whole number. (2)
- (d) Using your graphic display calculator, or otherwise, find the chi-squared statistic for Manuel's data. (3)
- (e)
  - (i) State the number of degrees of freedom available for this calculation.
  - (ii) State the critical value for Manuel's test.
  - (iii) State his conclusion.

(3)  
(Total 12 marks)



9. (a) For his Mathematical Studies project, Marty set out to discover if stress was related to the amount of time that students spent travelling to or from school. The results of one of his surveys are shown in the table below.

Travel time ( $t$ mins)	Number of students		
	high stress	moderate stress	low stress
$t \leq 15$	9	5	18
$15 < t \leq 30$	17	8	28
$30 < t$	18	6	7

He used a  $\chi^2$  test at the 5% level of significance to find out if there was any relationship between student stress and travel time.

- (i) Write down the null and alternative hypotheses for this test. (2)
- (ii) Write down the table of expected values. Give values to the nearest integer. (3)
- (iii) Show that there are 4 degrees of freedom. (1)
- (iv) Calculate the  $\chi^2$  statistic for this data. (2)

The  $\chi^2$  critical value for 4 degrees of freedom at the 5% level of significance is 9.488.

- (v) What conclusion can Marty draw from this test? Give a reason for your answer. (2)

- (b) Marty asked some of his classmates to rate their level of stress out of 10, with 10 being very high. He also asked them to measure the number of minutes it took them to get from home to school. A random selection of his results is listed below.

Travel time ( $x$ )	13	24	22	18	36	16	14	20	6	12
Stress rating ( $y$ )	3	7	5	4	8	8	4	8	2	6

- (i) Write down the value of the (linear) coefficient of correlation for this information. (1)
- (ii) Explain what a positive value for the coefficient of correlation indicates. (1)
- (iii) Write down the linear regression equation of  $y$  on  $x$  in the form  $y = ax + b$  (2)
- (iv) Use your equation in part (iii) to determine the stress rating for a student who takes three quarters of an hour to travel to school. (2)
- (v) Can your answer in part (iv) be considered reliable? Give a reason for your answer. (2)
- (Total 18 marks)**

10. The eye colour and gender of 500 students are noted and the results are indicated in the table below.

	Blue	Brown	Green
Male	18	152	50
Female	40	180	60

It is believed that eye colour is related to gender in a school in Banff. It is decided to test this hypothesis by using a  $\chi^2$  test at the 5% level of significance.

- (a) Write down the null hypothesis for this experiment. (1)
- (b) Show that the number of degrees of freedom is 2. (1)

(c) Write down the  $\chi^2$  critical value for the degrees of freedom. (1)

(d) Calculate the  $\chi^2$  test statistic for this data. (2)

(e) Does the evidence suggest that eye colour is related to gender in this school? Give a clear reason for your answer. (2)

**(Total 7 marks)**

11. In a competition the number of males and females taking part in different swimming races is given in the table of observed values below.

	Backstroke (100 m)	Freestyle (100 m)	Butterfly (100 m)	Breaststroke (100 m)	Relay (4 × 100 m)
Male	30	90	31	29	20
Female	28	63	20	37	12

The Swimming Committee decides to perform a  $\chi^2$  test at the 5% significance level in order to test if the number of entries for the various strokes is related to gender.

- (a) State the null hypothesis. (1)
- (b) Write down the number of degrees of freedom. (1)
- (c) Write down the critical value of  $\chi^2$ . (1)

The expected values are given in the table below:

	Backstroke (100 m)	Freestyle (100 m)	Butterfly (100 m)	Breaststroke (100 m)	Relay (4 × 100 m)
Male	32	$a$	28	37	18
Female	26	68	23	$b$	14

- (d) Calculate the values of  $a$  and  $b$ . (2)
- (e) Calculate the  $\chi^2$  value. (3)
- (f) State whether or not you accept your null hypothesis and give a reason for your answer. (2)
- (Total 10 marks)**

12. The veterinarian has gathered the following data about the weight of dogs and the weight of their puppies.

		Dog		Total
		Heavy	Light	
Puppy	Heavy	36	27	63
	Light	22	35	57
	Total	58	62	120

The veterinarian wishes to test the following hypotheses.

$H_0$ : A puppy's weight is independent of its parent's weight.

$H_1$ : A puppy's weight is related to the weight of its parent.

- (a) The table below sets out the elements required to calculate the  $\chi^2$  value for this data.

	$f_o$	$f_e$	$f_e - f_o$	$(f_e - f_o)^2$	$(f_e - f_o)^2 / f_e$
heavy/heavy	36	30.45	-5.55	30.8025	1.012
heavy/light	27	32.55	5.55	30.8025	0.946
light/heavy	22	27.55	5.55	30.8025	1.118
light/light	35	$a$	$b$	$c$	$d$

- (i) Write down the values of  $a$ ,  $b$ ,  $c$ , and  $d$ . (4)
- (ii) What is the value of  $\chi^2_{\text{calc}}$  for this data? (1)
- (iii) How many degrees of freedom exist for the contingency table? (1)
- (iv) Write down the critical value of  $\chi^2$  for the 5% significance level. (1)
- (b) Should  $H_0$  be accepted? Explain why. (2)

(Total 9 marks)

13. The following table of observed results gives the number of candidates taking a Mathematics examination classified by gender and grade obtained.

		Grade			Total
		5, 6 or 7	3 or 4	1 or 2	
Gender	Males	5000	3400	600	9000
	Females	6000	4000	1000	11000
	Total	11000	7400	1600	20000

The question posed is whether gender and grade obtained are independent.

- (a) **Show clearly** that the expected number of males achieving a grade of 5, 6 or 7 is 4950. (2)
- (b) A  $\chi^2$  test is set up.
- (i) State the Null hypothesis. (1)
- (ii) State the number of degrees of freedom. (1)
- (iii) The calculated  $\chi^2$  value at the 5% test level is 39.957.  
Write down the critical value of  $\chi^2$  at the 5% level of significance. (1)
- (iv) What can you say about gender and grade obtained? (1)

(Total 6 marks)

14. In the small town of *Joinville*, population 1000, an election was held. The results were as follows:

	Urban Voters	Rural Voters
Candidate A	295	226
Candidate B	313	166

In (a) to (c) below we will use a chi-squared test to decide whether the choice of candidate depends on where the voter lives.

Null Hypothesis  $H_0$ : The choice of candidate is independent of where the voter lives.

- (a) (i) Write down the alternative hypothesis.  
(ii) Use the information above to fill in  $a$  and  $b$  in the table below.

Cell	$f_0$	$f_e$	$f_0 - f_e$	$(f_0 - f_e)^2$
1	295	317	-22	484
2	226	204	22	484
3	313	291	22	484
4	166	$a$	$b$	484

(3)

- (b) (i) Calculate the chi-squared statistic.  
(ii) Write the number of degrees of freedom.  
(iii) At the 5% confidence level, state the chi-squared critical value.

(5)

- (c) (i) Hence, state your conclusion.  
(ii) Explain why you reached this conclusion.

(2)

(Total 10 marks)

15. A bag containing 60 sweets is opened. The bag contains sweets of the following colours.

Colour	Frequency
Red	18
Orange	17
Green	10
Purple	9
Blue	6

According to the manufacturer, the various colours should have the following percentages.

Colour	Percentage
Red	35%
Orange	25%
Green	20%
Purple	15%
Blue	5%

- (a) Calculate the expected number of sweets of each colour in a bag containing exactly 60 sweets.

(3)

Before you can perform the chi-squared test on this data, it is necessary to combine the data for one of the colours with that of another colour.

- (b) Which colour is this and why is this necessary?

(2)

- (c) Using the chi-squared test at the 5% significance level, investigate the hypothesis that the sweets in the packet may be regarded as a random sample. Remember to state the null hypothesis, the number of degrees of freedom and the critical value of chi-squared.

(7)

**(Total 12 marks)**



16. A survey was conducted in a company to determine whether position in upper management was independent of gender. The results of this survey are tabulated below.

	Managers	Junior executives	Senior executives	Totals
Male	95	130	75	<b>300</b>
Female	65	110	25	<b>200</b>
Totals	<b>160</b>	<b>240</b>	<b>100</b>	<b>500</b>

The table below shows the expected number of males and females at each level, if they were represented proportionally to the total numbers of males and females employed.

	Managers	Junior executives	Senior executives	Totals
Male	<i>a</i>	<i>c</i>	60	<b>300</b>
Female	<i>b</i>	<i>d</i>	40	<b>200</b>
Totals	<b>160</b>	<b>240</b>	<b>100</b>	<b>500</b>

- (a) (i) Show that the expected number of **Male Managers** (*a*) is 96.  
(ii) Hence find the values of *b*, *c* and *d*. (5)
- (b) (i) Write a suitable null hypothesis for this data.  
(ii) Write a suitable alternate hypothesis for this data. (2)
- (c) (i) Perform a chi-squared test of independence for this data to show the value of  $\chi^2$  is 12.8 to 3 significant figures.  
(ii) Calculate the number of degrees of freedom, and write down the critical value of  $\chi^2$  at the 5% significance level.  
(iii) What conclusion can be drawn regarding gender and position in upper management? (6)

(Total 13 marks)

17. For his Mathematical Studies Project a student gave his classmates a questionnaire to fill out. The results for the question on the gender of the student and specific subjects taken by the student are given in the table below, which is a  $2 \times 3$  contingency table of **observed** values.

	History	Biology	French	
Female	22	20	18	(60)
Male	20	11	9	(40)
	(42)	(31)	(27)	

The following is the table for the **expected** values.

	History	Biology	French
Female	$p$	18.6	16.2
Male	$q$	$r$	10.8

- (a) Calculate the values of  $p$ ,  $q$  and  $r$ .

(3)

The chi-squared test is used to determine if the choice of subject is independent of gender, at the 5% level of significance.

- (b) (i) State a suitable null hypothesis  $H_0$ .
- (ii) Show that the number of degrees of freedom is two.
- (iii) Write down the critical value of chi-squared at the 5% level of significance.

(3)

- (c) The calculated value of chi-squared is 1.78. Do you accept  $H_0$ ? Explain your answer.

(2)

(Total 8 marks)

18. Members of a certain club are required to register for one of three games, billiards, snooker or darts.

The number of club members of each gender choosing each game in a particular year is shown in the table below.

	Billiards	Snooker	Darts
Male	39	16	8
Female	21	14	17

- (a) Use a  $\chi^2$  (Chi-squared) test at the 5% significance level to test whether choice of games is independent of gender. State clearly the null and alternative hypotheses tested, the expected values, and the number of degrees of freedom used.

(13)

The following year the choice of games was widened and the figures for that year are as follows:

	Billiards	Snooker	Darts	Fencing
Male	4	15	8	10
Female	10	21	17	37

- (b) If the  $\chi^2$  test were applied to this new set of data,
- why would it be necessary to combine billiards with another game?
  - which other game would you combine with billiards and why?

(2)

A club member is to be selected at random.

- (c) What is the probability that the club member selected is a
- female who chose billiards or snooker?
  - male or female who chose darts or fencing?

(2)

(Total 17 marks)