



FS1 (Year 1) – Chi-Squared Contingency Tables Exam Questions (Edexcel)

Q1.

A doctor takes a random sample of 100 patients and measures their intake of saturated fats in their food and the level of cholesterol in their blood. The results are summarised in the table below.

Cholesterol level Intake of saturated fats	High	Low
	High	Low
High	12	8
Low	26	54

Using a 5% level of significance, test whether or not there is an association between cholesterol level and intake of saturated fats. State your hypotheses and show your working clearly.

(10)

(Total 10 marks)

(Q02 6691/01, June 2013)

Q2.

The Director of Studies at a large college believed that students' grades in Mathematics were independent of their grades in English. She examined the results of a random group of candidates who had studied both subjects and she recorded the number of candidates in each of the 6 categories shown.

	Maths grade A or B	Maths grade C or D	Maths grade E or U
English grade A or B	25	25	10
English grade C to U	15	30	15

(a) Stating your hypotheses clearly, test the Director's belief using a 10% level of significance. You must show each step of your working.

(9)

The Head of English suggested that the Director was losing accuracy by combining the English grades C to U in one row. He suggested that the Director should split the English grades into two rows, grades C or D and grades E or U as for Mathematics.

(b) State why this might lead to problems in performing the test.

(1)

(Total 10 marks)

(Q04 6691/01, June 2007)



Q3.

Students in a mixed sixth form college are classified as taking courses in either Arts, Science or Humanities. A random sample of students from the college gave the following results.

		Course		
		Arts	Science	Humanities
Gender	Boy	30	50	35
	Girl	40	20	42

Showing your working clearly, test, at the 1 % level of significance, whether or not there is an association between gender and the type of course taken. State your hypotheses clearly.

(11)

(Total 11 marks)

(Q03 6691/01, June 2008)

Q4.

A factory manufactures batches of an electronic component. Each component is manufactured in one of three shifts. A component may have one of two types of defect, D_1 or D_2 , at the end of the manufacturing process. A production manager believes that the type of defect is dependent upon the shift that manufactured the component. He examines 200 randomly selected defective components and classifies them by defect type and shift. The results are shown in the table below.

Shift \ Defect type	D_1	D_2
	First shift	45
Second shift	55	20
Third shift	50	12

Stating your hypotheses, test, at the 10% level of significance, whether or not there is evidence to support the manager's belief. Show your working clearly.

(10)

(Total 10 marks)

(Q02 6691/01, June 2011)



Q5.

A psychologist carries out a survey of the perceived body weight of 150 randomly chosen people. He asks them if they think they are underweight, about right or overweight. His results are summarised in the table below.

	Underweight	About right	Overweight
Male	20	22	30
Female	16	28	34

The psychologist calculates two of the expected frequencies, to 2 decimal places, for a test of independence between perceived body weight and gender. These results are shown in the table below.

	Underweight	About right	Overweight
Male	17.28		
Female	18.72		

(a) Complete the table of expected frequencies shown above.

(2)

(b) Test, at the 10% level of significance, whether or not perceived body weight is independent of gender. State your hypotheses clearly.

(7)

The psychologist now combines the male and female data to test whether or not body weight types are chosen equally.

(c) Find the smallest significance level, from the tables in the formula booklet, for which there is evidence of a preference.

(5)

(Total for question = 14 marks)

(Q03 6691/01, June 2017)

Q6.

Abram carried out a survey of two treatments for a plant fungus. The contingency table below shows the results of a survey of a random sample of 125 plants with the fungus.

		Treatment		
		No action	Plant sprayed once	Plant sprayed every day
Outcome	Plant died within a month	15	16	25
	Plant survived for 1 – 6 months	8	25	10
	Plant survived beyond 6 months	7	14	5

[Questions Continued On Next Page]



Abram calculates expected frequencies to carry out a suitable test. Seven of these are given in the partly-completed table below.

		Treatment		
		No action	Plant sprayed once	Plant sprayed every day
Outcome	Plant died within a month			17.92
	Plant survived for 1 – 6 months	10.32	18.92	13.76
	Plant survived beyond 6 months	6.24	11.44	8.32

$$\sum \frac{(O - E)^2}{E}$$

The value of $\sum \frac{(O - E)^2}{E}$ for the 7 given values is 8.29

Test at the 2.5% level of significance, whether or not there is an association between the treatment of the plants and their survival. State your hypotheses and conclusion clearly.

(7)

(Total for question = 7 marks)

(Q04 8FM0/23, June 2018)

Q7.

A university foreign language department carried out a survey of prospective students to find out which of three languages they were most interested in studying.

A random sample of 150 prospective students gave the following results.

		Language		
		French	Spanish	Mandarin
Gender	Male	23	22	20
	Female	38	32	15

A test is carried out at the 1% level of significance to determine whether or not there is an association between gender and choice of language.

(a) State the null hypothesis for this test.

(1)

(b) Show that the expected frequency for females choosing Spanish is 30.6

(1)

(c) Calculate the test statistic for this test, stating the expected frequencies you have used.

(3)

(d) State whether or not the null hypothesis is rejected. Justify your answer.

(2)

(e) Explain whether or not the null hypothesis would be rejected if the test was carried out at the 10% level of significance.

(1)

(Total for question = 8 marks)

(Q01 8FM0/2G/sA, Specimen papers)



Q8.

		Treatment	
		Applied	Not applied
Number of years free from this disease	$y < 1$	15	25
	$1 \leq y < 2$	35	61
	$2 \leq y$	124	140

Stuart is investigating a treatment for a disease that affects fruit trees. He has 400 fruit trees and applies the treatment to a random sample of these trees. The remainder of the trees have no treatment. He records the number of years, y , that each fruit tree remains free from this disease.

The results are summarised in the table (left).

The data are to be used to determine whether or not there is an association between the application of the treatment and the number of years that a fruit tree remains free from this disease.

- (i) Applied and $y < 1$
 (ii) Not applied and $1 \leq y < 2$

(2)

The value of $\sum \frac{(O - E)^2}{E}$ for the **other four** classes is 2.642 to 3 decimal places.

(b) Test, at the 5% level of significance, whether or not there is an association between the application of the treatment and the number of years a fruit tree remains free from this disease.

You should state your hypotheses, test statistic, critical value and conclusion clearly.

(5)

(Total for question = 7 marks)

(Q01 8FM0/23, June 2022)

Q9.

Charlie carried out a survey on the main type of investment people have. The contingency table below shows the results of a survey of a random sample of people.

		Main type of investment		
		Bonds	Cash	Stocks
Age	25–44	a	$b - e$	e
	45–75	c	$d - 59$	59

(a) Find an expression, in terms of a , b , c and d , for the difference between the observed and the expected value ($O - E$) for the group whose main type of investment is Bonds and are aged 45 – 75
 Express your answer as a single fraction in its simplest form.

(4)

Given that $\sum \frac{(O - E)^2}{E} = 9.62$ for this information,

(b) test, at the 5% level of significance, whether or not there is evidence of an association between the age of a person and the main type of investment they have.

You should state your hypotheses, critical value and conclusion clearly.

You may assume that no cells need to be combined.

(3)

(Total for question = 7 marks)

(Q04 8FM0/23, Oct 2021)