Bio Factsheet

Number 79

The Chi-SquaredTest for Goodness of Fit

The chi-squared (χ^2) test is widely used within project work. This Factsheet will tell you when and how to use it. Questions using the chi-squared test may occur on exam papers, but you will not have to remember the formula - it will be given to you.

Hypotheses

The purpose of all statistical tests is to choose between two *hypotheses* - for example: "the leaves are smaller at the top of the tree than at the bottom" and "the leaves are not smaller at the top of the tree than at the bottom". The *null hypothesis* (H_0) always has to be the "boring case" - that there's no difference between things. In the above example, it would be "the leaves are not smaller at the top of the tree" - or equivalently, "the leaves are the same size at the top and bottom of the tree". It can **never** be "the leaves are bigger at the top of the tree". The other hypothesis is called the *alternative hypothesis* (H_1). In our leaves example, it would be "the leaves are smaller at the top of the tree".

When we carry out a test, we always start out assuming that the null hypothesis is true, and we only change our minds if we have enough evidence - it's like a trial, when you are assumed to be innocent unless there's enough evidence to show that you aren't! Table 1 shows some possible investigations using chi-squared and the corresponding null hypotheses.

Table 1. Investigations using chi-squared

Investigation	Null Hypothesis (H_0)	What to Measure
How the concentration of nitrate in water affects seed germination.	The concentration of nitrate in water has no effect on seed germination	The number of seeds germinating within a set of time period when watered with solutions containing different concentrations of nitrate.
How a particular type of pollution affects a particular organism	The level of pollution has no effect on the numbers of the organism	The numbers of the organism obtained from a set area in two sites which are similar, except for one being polluted and one not polluted
Are the predictions of genetics accurate?	The numbers of organisms in each category are in accordance with the predictions of genetics	The numbers of organisms in each category

What is chi-squared?

Chi-squared goodness of fit is used to test whether the actual results of an experiment fit in with what we'd expect if the null hypothesis were true. To see how this works, imagine testing a normal coin to see whether heads and tails were equally likely. Our null hypothesis is that they are equally likely - we have no reason to believe otherwise before we do an experiment, and we always choose "the boring case" for the null hypothesis. If we tossed our coin 600 times, we'd expect to get about half of each - around 300 heads and 300 tails.

If we then actually tossed the coin 600 times and got over 500 heads, we'd feel that this was a long way off from our predictions - so we'd probably decide that the coin was weighted. However, if we got 305 heads and 295 tails, we'd probably feel this was close enough, and decide the coin was OK. The chi-squared test lets us decide on an accurate basis what counts as "close enough".

Obviously, we can never be absolutely certain that our decision is correct - we could get 500 heads by chance even if the coin wasn't weighted. We can decide how far off the results have to be by carrying out the test at different *significance levels.* The smaller the significance level we use, the "further off" the results need to be for us to reject the null hypothesis. Using a smaller significance level is like requiring the evidence to be more convincing. Statistical tests in biology are usually carried out at the 5% significance level.

Exam Hint: *- Marks are only awarded for an appropriate use of statistics. Decide exactly what your hypotheses are and what test you are going to use before collecting your data.*

When can chi-squared be used?

You can only use chi-squared when you have *frequencies* - that is, numbers of items in particular categories. You cannot use it to compare measurements or other figures directly. For example, if you were doing an experiment on seed germination, you could use chi-squared to compare the numbers of seeds germinating within a week in each of four different solutions. However, you could not use it to compare the heights of four different seedlings.

In order that the test be valid, it is also important that you should expect **at least five items in each category**. Sometimes it may be necessary to combine categories in order to achieve this - for example, if you were researching public attitudes by using a questionnaire, you might need to combine the responses "not very concerned" and "not at all concerned".

Exam Hint: *- Do not try to be too ambitious in your investigation. Testing one simple, easily measurable hypothesis involving only one variable successfully will gain more marks than an attempt at investigating a situation affected by many variables.*

Worked Example

A student decides to investigate whether pollution levels affect the incidence of asthma. They obtain a sample of 50 sixth formers from their own school, which is situated in a large, polluted city and 50 sixth formers from another school, which is situated in a small, relatively unpolluted town. Each sixth former is asked whether or not s/he suffers from asthma.

The student obtains the following results: City: 16 with asthma Town: 8 with asthma

Step 1: Write down the **hypotheses**

Hypotheses are:
H_0 (null hypothesis): Pollution has no effect on incidence of asthma
H_1 (alternative hypothesis): Pollution has some effect on the incidence of asthma

Step 2: Work out the **expected frequencies**

We do this by adding up all the people with asthma, and dividing by the number of different categories we're looking at - which is two (city and town).
So the expected frequencies are $(16 + 8) \div 2 = 12$ in each area

Step 3: Put your observed frequencies (from the actual experiment) and the expected frequencies (from step 2) in a **table.**

	City	Town
Observed (O)	16	8
Expected (E)	12	12

Exam Hint: *- Don't worry if your expected frequencies are not whole numbers. They don't have to be!* ***Do not*** *round them to the nearest whole number - this will make your test less accurate.*

Step 4: For each of your categories, work out $\frac{(\mathbf{O - E})^2}{\mathbf{E}}$

For the city: $\frac{(16 - 12)^2}{12} = 1.33$ For the town: $\frac{(8 - 12)^2}{12} = 1.33$

Step 5: **Add up** all these values.
This gives the **chi-squared value**

chi-squared value = 1.33 + 1.33 = 2.66

Step 6: Work out the **degrees of freedom.**
This is one less than the number of categories

Degrees of freedom = 2 – 1 = **1**

Exam Hint: *- Don't worry about what degrees of freedom means! Unless you want to study Statistics as a subject, you don't need to know!*

Step 7: Get a chi-squared table and **look up the value** for the appropriate significance level (usually 5%) and the degrees of freedom. In the table 5% is shown as 0.05.

We look for the 5% level for one degree of freedom.
This is **3.84** (Table 2)

Table 2. Chi-squared tables

df	0.10	**0.05**	0.025	0.01	0.005
1	2.71	**3.84**	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.23	14.86

Step 8: **Make a decision** - if your chi-squared value is **bigger** than the one from the tables, you can **reject** the null hypothesis. Otherwise you have to accept it.

Our value is smaller than the value from the tables, so we accept the null hypothesis - there is no significant difference in the amount of asthma between the city and the town.

Points to note

1. This investigation needs care in sampling technique!
- The sixth formers need to live in the area, not just go to school in it
- They may not have lived there for long
- School sixth formers may not be representative of the population as a whole
- How does the student know what the pollution levels are - are they just assuming it?
- Pollution levels are not the same throughout a city or town.

2. This test is not telling us that the null hypothesis is definitely correct - it is telling us that we haven't got enough evidence to reject it.

3. To improve the chance of getting a significant result - in other words, rejecting the null hypothesis - a larger sample usually helps! For example, if the student had taken a sample of 100 from each school and found the city and town had 32 and 16 respectively with asthma, s/he would have been able to reject the null hypothesis (check this calculation!)

Worked Example

A student decides to investigate the results of crossing plants with red and yellow flowers. The student knows that the allele for red flowers is dominant. S/he carries out the cross, and obtains 18 plants, all of which have red flowers.

a) Give the genotype of the 18 red-flowered plants produced. **1**

b) Explain what, if anything, can be deduced about the genotype of the parent red-flowered plant. **2**

The student then crosses two of the offspring red-flowered plants and obtains 17 red-flowered plants and 3 yellow-flowered plants.

c) i) Find the ratio of red-flowered to yellow-flowered plants that would be expected. **2**

ii) Carry out a chi-squared test at the 5% significance level to determine whether the results are in accordance with your predictions. **10**

Answer and mark Scheme

a) It must be Rr (since the yellow parent would be rr, and any rr offspring would be yellow); **1**

b) Probably RR; since no yellow-flowered offspring are produced, (but this is not certain); **2**

c) i) Rr crossed with Rr produces RR, Rr, rR and rr in equal proportion; Since the first three all are red-flowered, we would expect red-flowered:yellow-flowered to by 3:1; **2**

ii)

Step 1: Write down the **hypotheses**

Hypotheses are:

H_0: Results obtained are not significantly different from 3:1 ratio

H_1: Results obtained are significantly different from 3:1 ratio; **1**

Step 2: Work out the **expected frequencies**

We do this using

$$\text{Expected no. of individuals in a category} = \frac{\text{Total no. of individuals} \times \text{Ratio number for that category}}{\text{Total of all the numbers in the ratio}}$$

So expected for red $= \frac{20 \times 3}{1 + 3} = 15$;

Expected for yellow $= \frac{20 \times 1}{1 + 3} = 5$; **2**

Exam Hint: *- Check that your expected frequencies add up to the total number of individuals - 15 + 5 = 20*

Step 3: Put your observed frequencies (from the actual experiment) and the expected frequencies (from step 2) in a **table.**

	Red	Yellow
O	17	3
E	15	5

1

Step 4: For each of your categories, work out $\frac{(\mathbf{O - E})^2}{\mathbf{E}}$

Red: $\frac{(17-15)^2}{15} = 0.2667$; Yellow: $\frac{(3-5)^2}{5} = 0.8$; **2**

Step 5: **Add up** all these values. This gives the **chi-squared value**

chi-squared value = 0.2667 + 0.8 =1.0667; **1**

Step 6: Work out the **degrees of freedom.** This is one less than the number of categories

Degrees of freedom = 2 – 1 = **1**; **1**

Step 7: Get a chi-squared table and **look up the value** for the appropriate significance level (usually 5%) and the degrees of freedom.

We look for the 5% level for one degree of freedom; This is **3.84** (Table 2 overleaf) **1**

Step 8: **Make a decision** - if your chi-squared value is **bigger** than the one from the tables, you can **reject** the null hypothesis. Otherwise you have to accept it.

Our value is smaller than the value from the tables, so we accept the null hypothesis - the results are not significantly different from the 3:1 ratio; **1**

Total 15

Acknowledgments: This Factsheet was researched and written by Cath Brown. *Curriculum Press, Unit 305B The Big Peg, 120 Vyse Street, Birmingham B18 6NF*
 ISSN 1351-5136