

## Chi Squared Test for Ecology (Using TiNspire)

The chi squared ( $\chi^2$ ) test can determine if two variables are **independent** of each other. It allows us to investigate relationships between species or the distribution of species within a habitat for example.

The chi squared ( $\chi^2$ ) test we perform is called a  $\chi^2$  **2-way test**. (Note that this is different from the “*goodness of fit*” test used in genetics.). The purpose of the test is to tell us if our sample significantly differs from the frequencies we would expect *if the variables are independent of each other*.

### Null and Alternate Hypotheses

Every  $\chi^2$  test should start with the statement of the **Null ( $H_0$ )** and **Alternate ( $H_1$ )** hypotheses.

The **null hypothesis** always states that the two variables are **independent** (and therefore that the presence or absence of one variable does not effect the presence or absence of the other).

The **alternate hypothesis** always states that the two variables are **not independent** (and therefore that one variable effects the other).

### Contingency Tables for Observed Results

The results from sampling data (the “**observed**” data) must be displayed in a suitable table so that each variable is listed either horizontally or vertically. This is called a contingency table. A contingency table has 2 or more rows and 2 or more columns of data (not including the totals).

### Expected Values

The expected results are calculated by assuming that the two variables are completely independent and calculating probabilities for each value in the contingency table. The chi square test compares the observed values with the expected values – the closer they are, the higher the likelihood that the variables are independent of each other.

### Degrees of Freedom

The  $\chi^2$  distribution is dependent on the number of degrees of freedom (df).

Degrees of freedom are calculated from the number of rows and columns (not including the totals) in the contingency table.

$$\text{Degrees of freedom} = (\text{number of rows} - 1)(\text{number of columns} - 1)$$

## Example question

A zoo veterinarian has gathered the following data about the size of lionesses and the weight of their cubs.

		Lioness		Total
		Large	Small	
Cub	Heavy	36	27	63
	Light	22	35	57
Total		58	62	120

The veterinarian wishes to test the following hypotheses.

$H_0$ : A cub's weight is independent of its mother's size.

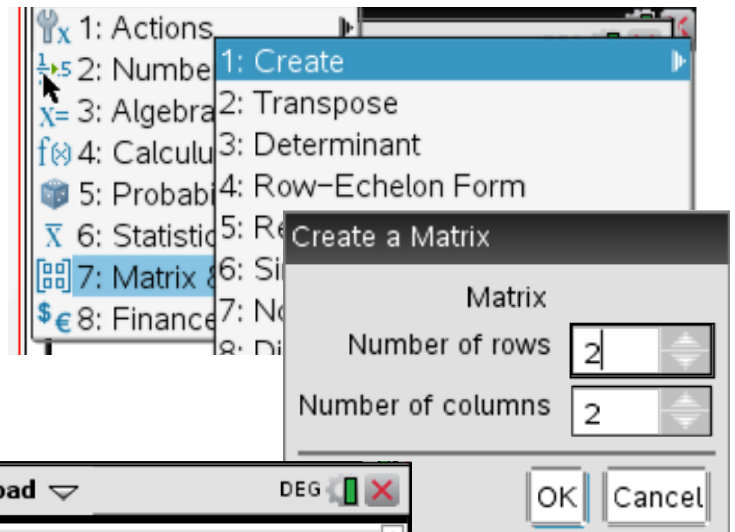
$H_1$ : A cub's weight is related to the size of its mother.

### Using the Calculator:

#### Entering observed values

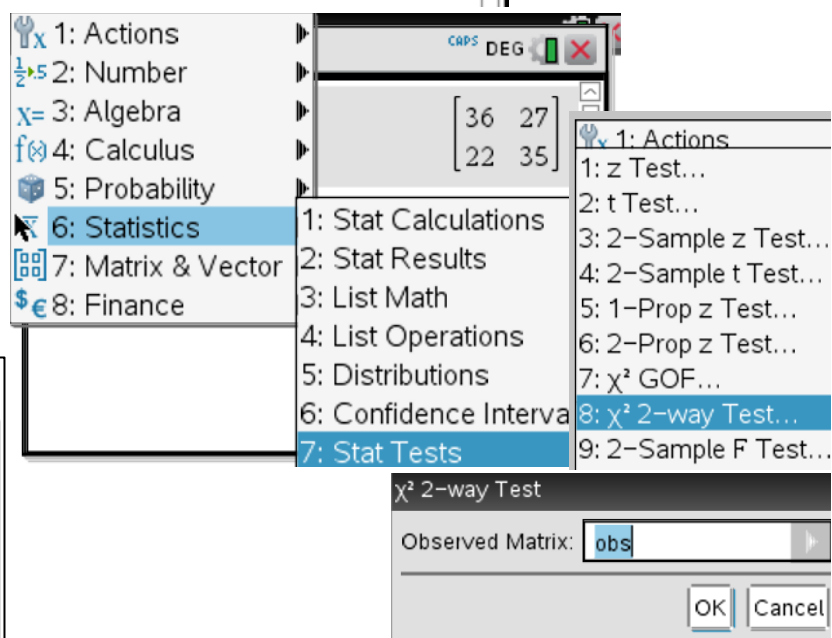
Open a calculate scratchpad

- Press MENU → 7: Matrix & Vector → 1: Create → 1: Matrix
- Enter the number of rows & columns, OK
- A matrix is created for you to enter the **observed values** into.  
Use TAB to move around the matrix.
- Make sure your cursor is outside the matrix, press CTRL VAR and type "obs" press ENTER



#### Finding $\chi^2$ and P values

- Press MENU → 6: Statistics → 7: Stat Tests → 8:  $\chi^2$  2-way Test, ENTER
- Choose "obs" from menu, OK
- Information Given:  
The results are the value for  $\chi^2$ , the p-value and the degrees of freedom.



$\chi^2$ 2way obs: stat.results	
"Title"	" $\chi^2$ 2-way Test"
" $\chi^2$ "	4.1218731
"PVal"	0.04233218
"df"	1.
"ExpMatrix"	"[...]"
"CompMatrix"	"[...]"

## Finding Expected Values

NOTE: You **MUST** calculate  $\chi^2$  first for this to work!

- Press VAR → stat.expmatrix, ENTER
- Information Given:  
The expected values are in the *same position* as the observed values

$\chi^2$  2-way obs: stat results

obs	" $\chi^2$ 2-way Test"
stat.compmatrix	4.1218731
stat.df	0.04233218
stat.expmatrix	1.
stat.pval	"[...]"
stat.results	"[...]"
stat.stat	
stat.values	

stat.ExpMatrix

30.45	32.55
27.55	29.45

## $\chi^2$ values, p-values and critical values

We compare either the  $\chi^2$  value to the critical value **or** the p-value (found with GDC) to 0.05 to determine if the null hypothesis is accepted or rejected.

**Accepting the null hypothesis** means the two variables are independent.

**Rejecting the null hypothesis** means the two variables are not independent (& the presence of one of the variables affects the other).

We always test at the **5% significance level**. This means we are 95% confident that the differences between observed and expected values are only due to chance.

The **p-value** is calculated on your GDC. You can ALWAYS accept or reject using this (rather than the critical value).

If the **p-value is more than** the significance level ( 0.05 ) we **accept the null hypothesis**.

If the **p-value is less than** the significance level ( 0.05 ) we **reject the null hypothesis**.

In an exam, you may be given a **critical value** (at 95% significance level). This is used to compare with the  $\chi^2$  value.

If the  $\chi^2$  value is **less than** the **critical value**, we **accept the null hypothesis**.

If the  $\chi^2$  value is **more than** the **critical value**, we **reject the null hypothesis**.

In our example problem above, the critical value for 1 degree of freedom = 3.841 at 5% significance level.

The  $\chi^2$  value = 4.12 which is > 3.841, so the  $H_0$  is rejected.

Alternatively, the p-value is 0.0423 which is < 0.05, so once again the  $H_0$  is rejected.

## Another Example Question<sup>1</sup>

Moorlands are upland areas with acidic and low-nutrient soils, where heather plants dominate. Heather plants are flowering plants with long woody stems which grow in dense clumps. The presence of two species of heather; bell heather (*Erica onerea*) and common heather (*Calluna vulgaris*), was investigated to see if they are associated together.

The presence or absence of common or bell heather was recorded in 200 random quadrat samples.

Results:

Species	Frequency
Bell heather only	31
Common heather only	45
Both species	89
Neither species	35

1. Construct a contingency table of observed values.
2. Calculate the number of degrees of freedom.
3. Calculate the chi-squared value.
4. State the two hypotheses,  $H_0$  and  $H_1$  for this example.
4. The critical value for a significance level of 5% is 3.841.

Decide if the  $H_0$  should be accepted or rejected, giving a reason for your decision.

1. Contingency table = observed results.

		Bell Heather	
		present	absent
Common Heather	present		
	absent		

2. Degrees of freedom =
3. Chi-square value =
4.  $H_0$  :
- $H_1$  :
5. Chi-square value =