

## 4.11 Statistical tests


### 4.11.1 $\chi^2$ test for independence

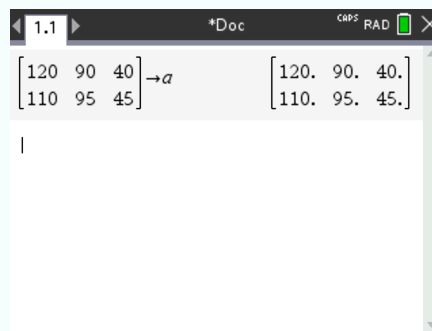
Consider the following set of data:

	Action	Horror	Comedy	Total
color-blind	120	90	40	250
non color-blind	110	95	45	250
Total	230	185	85	500



To be able to do a  $\chi^2$  test, you first need to put the data in a matrix.

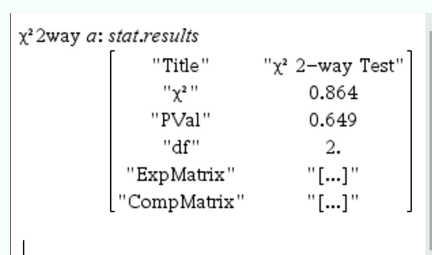
#### Enter the data

- Press , select Matrix & Vector > Create > Matrix. Set the matrix amount of rows and columns (here:  $2 \times 3$ ), and enter the data.
- Store the matrix as A. To do that press  then  and enter A:



#### Do the test

- Press , select Statistics > Stat Tests >  $\chi^2$  2-way Test. Set matrix [A] as Observed. Press , these results should be displayed:



### 4.11.2 $\chi^2$ goodness of fit test

Consider a person counting the amount of cyclists he sees passing by his street each day:

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
50	60	42	48	52	58	61

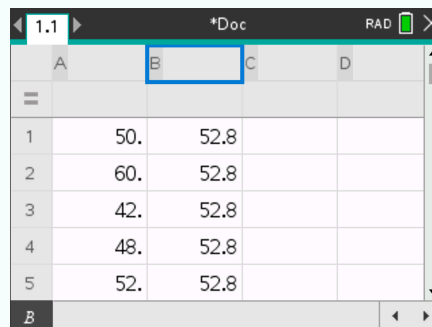
The null and alternative hypothesis are

$H_0$  : An equal amount of cyclists pass by his street each day.

$H_1$  : A different amount of cyclists pass by his street each day.

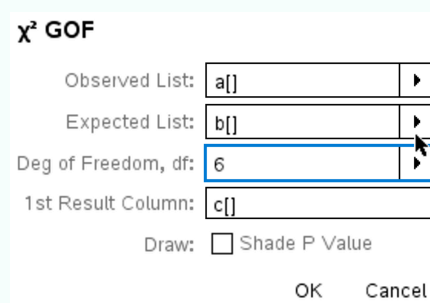
We want to know at a significance level of 0.05 if he must accept null hypothesis.

- ① Create a new document and select Add List & Spreadsheet.
- ② Fill column A with the amount of cyclist each day. Fill column B with the average amount of cyclists (here: 52.8).



	A	B	C	D
1	50.	52.8		
2	60.	52.8		
3	42.	52.8		
4	48.	52.8		
5	52.	52.8		

- ③ Press  , select Statistics > Stat Tests >  $\chi^2$  GOF. Fill the parameters as follows:



**$\chi^2$  GOF**

Observed List:

Expected List:

Deg of Freedom, df:

1st Result Column:

Draw:  Shade P Value

OK Cancel

Press  . These results should be displayed:

Title	$\chi^2$ GOF
$\chi^2$	5.57
PVal	0.473
df	6.
CompLis...	{0.14848...

The results should be  $\chi^2 = 5.57$  (for the critical value) and  $p = 0.473$  (for the significance level), rounded.

We must then accept the null hypothesis.

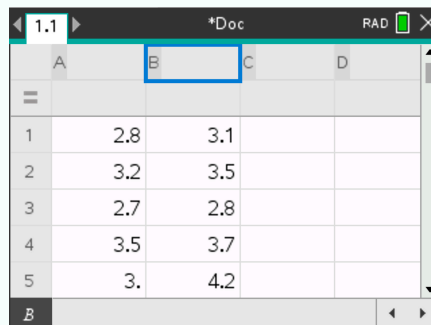
#### 4.11.3 The student's t-test

Consider the following data:


$x_1$	2.8	3.2	2.7	3.5	3.0	2.9	4.1	3.9	
$x_2$	3.1	3.5	2.8	3.7	4.2	2.6	3.2	2.9	3.8

You want to test whether the  $x_1$  data is on average a than  $x_2$  ( $\mu_1 > \mu_2$ ), at a significance level of 10%

- ① Create a new document and select Add List & Spreadsheet. Fill column A with  $x_1$  values and column B with  $x_2$  values.



	A	B	C	D
=				
1	2.8	3.1		
2	3.2	3.5		
3	2.7	2.8		
4	3.5	3.7		
5	3.	4.2		

- ② Press , select Statistics > Stat Tests > 2-Sample t Test. Select Data as data input and fill the parameters as follows:

**2-Sample t Test**

List 1: a[]

List 2: b[]


Frequency 1: 1

Frequency 2: 1

Alternate Hyp:  $H_a: \mu_1 > \mu_2$

Pooled: No

OK Cancel

Press  . These results should be displayed:

Title	2-Samp...
Alternate...	$\mu_1 > \mu_2$
t	-0.191
PVal	0.575
df	14.8

The  $t$ -value should be  $t = -0.191$  and the  $p$ -value should be  $p = 0.575$  (rounded). Therefore we must accept the null hypothesis (we **cannot** infer that  $\mu_1 > \mu_2$ ).