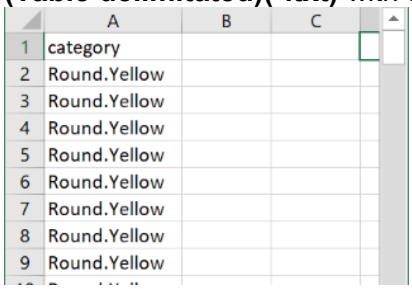
R HELP SHEET: One-Way Chi-Square

CONTENTS

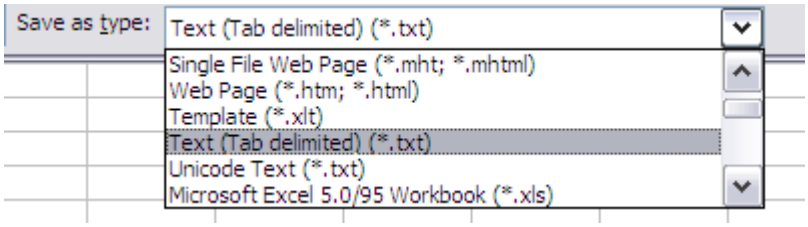
1. Creating a tab delimited data file using Excel

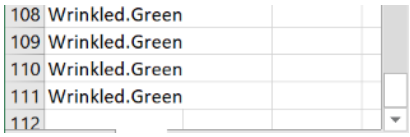
NOTE: If your data are already frequencies you can skip this and go to step 2 of conducting a one-way chi-square test.

Open Excel and type your observations (in this example *Round.Yellow* etc) in to a single column with an appropriate heading at the top of (in this example *category*), then save the file as a **Text (Table delimitated)(*.txt)** with an appropriate name (e.g., *RData_Peas*) file using **Save as type.**



	A
1	category
2	Round.Yellow
3	Round.Yellow
4	Round.Yellow
5	Round.Yellow
6	Round.Yellow
7	Round.Yellow
8	Round.Yellow
9	Round.Yellow





108	Wrinkled.Green
109	Wrinkled.Green
110	Wrinkled.Green
111	Wrinkled.Green
112	

2. Conducting a one-way chi-square test

The text in green after the hash (#) sign is just **notes** to help you remember what's in the output: it does not get R to actually "do" anything. The text in blue is **R code** with stars representing words that are specific to the example: you need to replace this with text specific to your data as shown in the output in section 3.

Step 1: If your data are raw observations, you will need to perform this step first, otherwise go straight to straight to step 2.

Open an **R-Editor** window by selecting **File** then **New script**.
Type in (or copy and paste) the notes and code below.
Replace the stars with appropriate text as indicated in notes.
Highlight everything and press **Ctrl R**.

```
#Importing data from tab delimited file
#(replace stars with an appropriate object name e.g.,peas)
****<-read.table(file.choose(),header=TRUE)
attach(****)
names(****)

#Calculating observed frequencies
#(replace stars with appropriate text e.g., category, category)
tapply(********,********,length)
```


Acknowledgement: Alvin Helden & Will Hoppitt for help with R code

Step 2: Once you have the frequencies for your data (or if that's the form they were already in) then to get R to conduct a one-way chi-square test

Type in (or copy and paste) the notes and code below.
Replace the stars with appropriate numbers as indicated in notes.
Highlight and press **Ctrl R**.

```
# Conducting a one-way chi-square replace stars with appropriate
# observed frequencies e.g., 26,31,26,27 obtained from step 1 and
# expected ratio as proportions e.g.,1/4,1/4,1/4,1/4
chisq.test(c(**,**,**,**),p=c((*/*),(*/*),(*/*),(*/*)))
```

3. Identifying the key elements of the output

Following the instructions above will produce the following output in the **R Console** window: the **key elements** are annotated in orange.

```
> #Importing data from tab delimited file (replace stars with an appropriate object name e.g.,peas)
> peas<-read.table(file.choose(),header=TRUE)
> attach(peas)
> names(peas)
[1] "form"  "category"
>
> #Calculating observed frequencies (replace stars with approriate text eg., category, category)
> tapply(category,category,length)
   Round.Green   Round.Yellow  Wrinkled.Green Wrinkled.Yellow
            26            31            26            27
>
> #Conducting a one-way chi-square replace stars with appropriate
> #observed frequencies e.g 26,31,26,27
> #and expected ratios as proportions and 1/4,1/4,1/4,1/4
> chisq.test(c(26,31,26,27),p=c((1/4),(1/4),(1/4),(1/4)))

        Chi-squared test for given probabilities

data:  c(26, 31, 26, 27)
X-squared = 0.6182, df = 3, p-value = 0.8923
```

NOTE: Use these frequencies generated by step 1 in the code in step 2

Statistic Degrees of Freedom P Value

In summary the key information from the test is

one-way chi-square: X^2_3 = 0.618, *N* = 110, *P* = 0.892

4. Additional notes

a. To find total sample size (N) use the following code:

```
#To find sample size
length(********)
```

For example: `length(category)`

b. Code for if there were only two categories would look like this:

```
chisq.test(c(**,**),p=c((*/**),(*/**)))
```

For example with observed frequencies 77 and 23 and expected ratio 3:1 as proportions 3/4,1/4

```
chisq.test(c(77,23),p=c((3/4),(1/4)))
```

c. Very small values of P will appear in scientific notation. You can run the following code before running the test to make it appear in decimal:

```
options(scipen=999)
```

And to turn scientific notation on again use the code:

```
options(scipen=0)
```


Acknowledgement: Alvin Helden & Will Hoppitt for help with R code