CHI-SQUARE TEST OF INDEPENDENCE: HINTS

YOU WILL FIND IN THIS FILE GENERAL APPROACH ON HOW TO APPLY CHI-SQUARE TEST. ADAPT THE METHOD TO YOUR DATA AND YOUR REQUESTS!

CHI-SQUARE TEST OF INDEPENDENCE: THEORETICAL BACKGROUND

In research, when we are interested in investigating the relationship between qualitative variable, the appropriate test is sometimes the chi-square test of independence. The steps necessary to be follow in order to carry out this test are:

- Calculate the observed frequencies for each cell of the cross tabulation (the Pivot Table can be used here);
- Calculate the expected frequencies for each cell of the cross tabulation (there is not a predefined formula to calculate expected frequencies);
- Compute the value of chi-square parameter;
- Evaluate the significance of obtained chi-square parameter.

The formula for the chi-square test of independence is give by:

$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ where O_i is an observed cell frequency and E_i is an expected cell frequency.

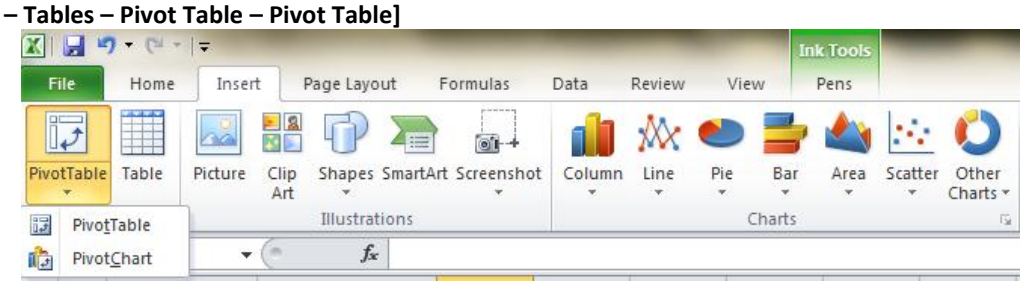
Assumptions for the Chi-Square Test of Independence

1. Subjects are randomly and independently sampled from the population of interest
2. Measurements are obtained from a single sample
3. Variables included in the analysis are measured on a qualitative scale
4. Expected cell frequencies are greater than or equal to five.

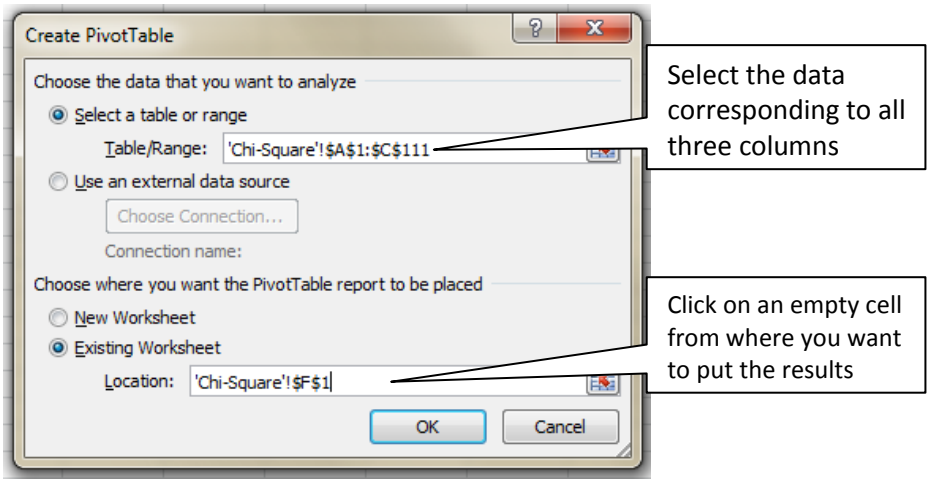
APPLYING CHI-SQUARE TEST: MICROSOFT EXCEL

To create the observed contingency table:

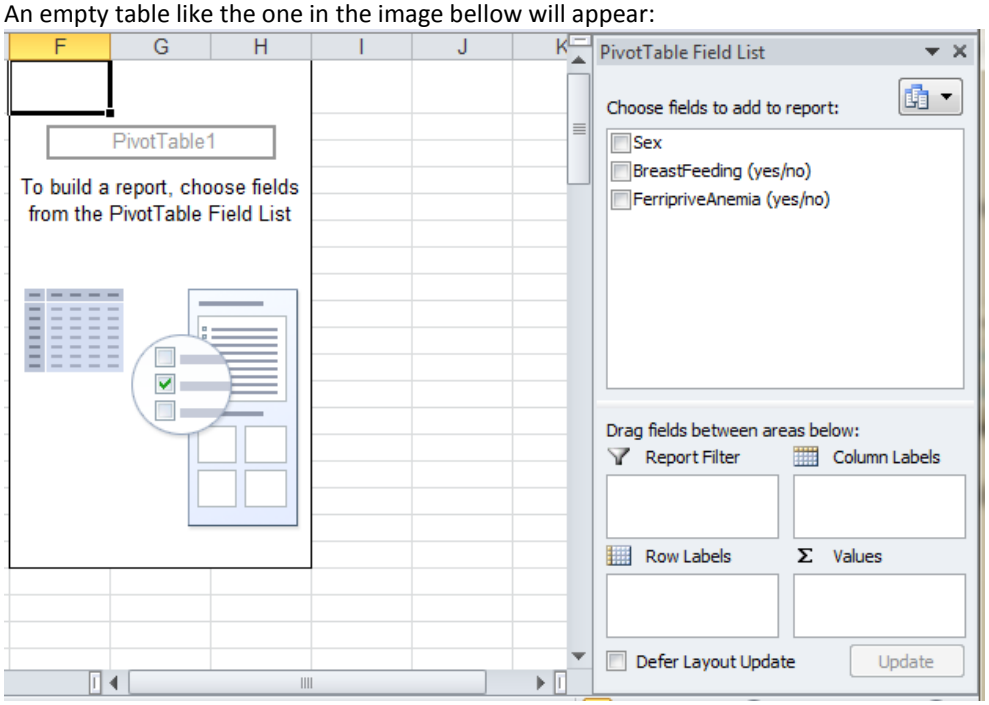
- **[Insert – Tables – Pivot Table – Pivot Table]**



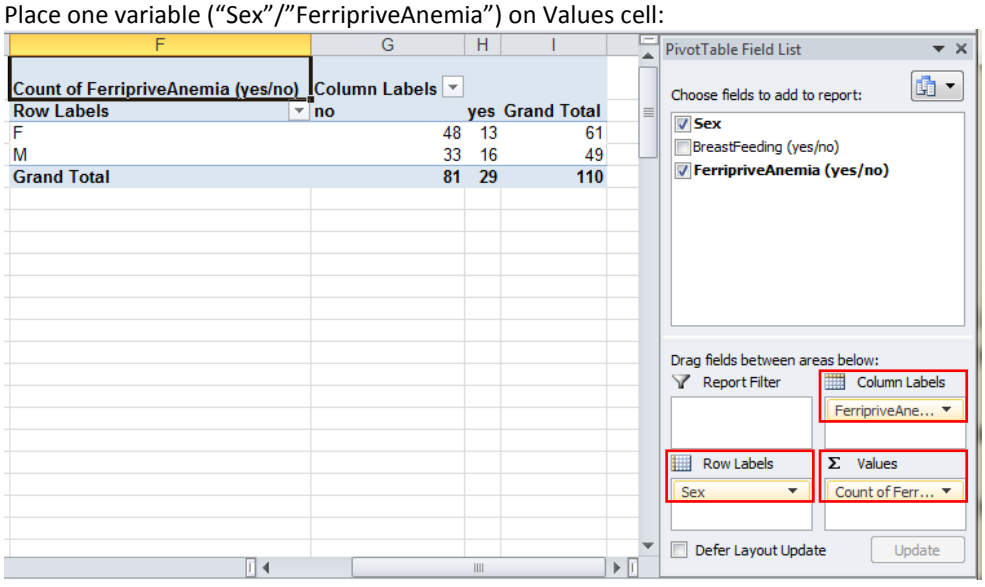
- Select data corresponding to all three variables as well as the cell where you want the results:



- An empty table like the one in the image bellow will appear:



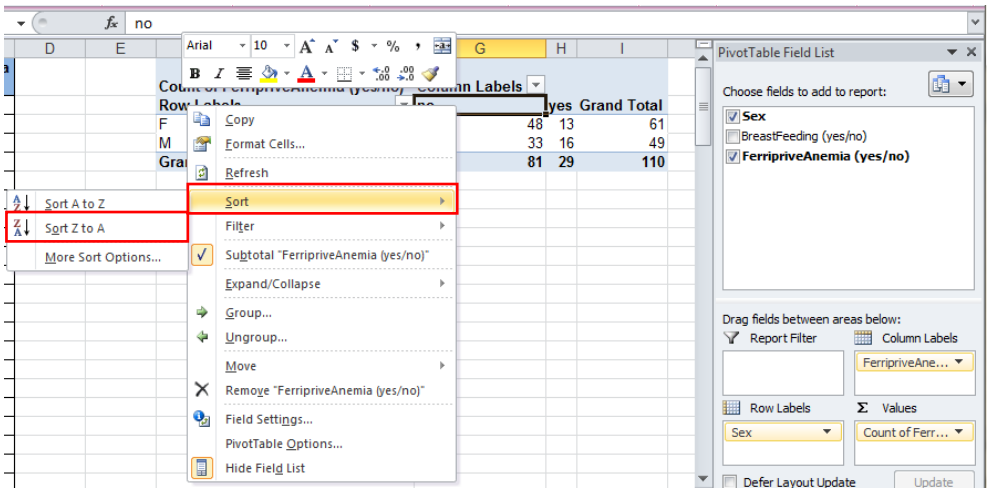
- Place the "*BreastFeeding*"/"*Sex*" variable on rows and "*FerripriveAnemia*" on columns. Place one variable ("Sex"/"FerripriveAnemia") on Values cell:



- Rearrange the table as in the image bellow:

Count of FerripriveAnemia (yes/no)	Column Labels		
Row Labels	yes	no	Grand Total
F	13	48	61
M	16	33	49
Grand Total	**29**	**81**	**110**

Select the cell 'no' + right click and [Sort – Sort Z to A]



- Copy and paste the observed contingency table and change the table as in the image bellow:

Observed table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	13	48	61
M	16	33	49
Grand Total	**29**	**81**	**110**

To create the expected contingency table:

- Copy the GrandTotal structure of the observed table:

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F			61
M			49
Grand Total	**29**	**81**	**110**

- Apply the following formula to calculate the expected cell frequencies:

$$E_i = \frac{(\text{Column Grand Total}) \cdot (\text{Row Grand Total})}{\text{Overall Grand Total}}$$

- To apply the formula:

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	=I15*G17/I17		61
M			49
Grand Total	**29**	**81**	**110**

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	16	=I15*H17/I17	61
M			49
Grand Total	**29**	**81**	**110**

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	16	45	61
M	=I16*G17/I17		49
Grand Total	**29**	**81**	**110**

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	16	45	61
M	13	=I16*H17/I17	49
Grand Total	**29**	**81**	**110**

  - The numbers in Expected table should be:

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	16	45	61
M	13	36	49
Grand Total	**29**	**81**	**110**

To compute chi-square statistics:

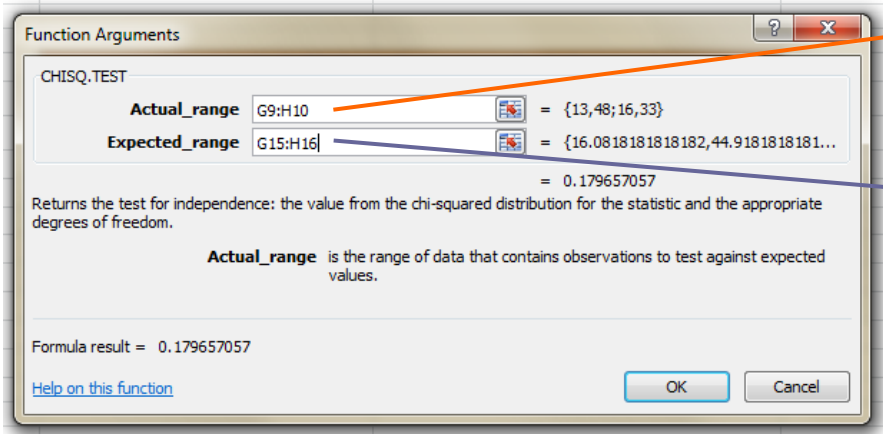
Chi-square test of independence: sex vs ferriprive anemia	
H_0 hypothesis	Sex and Feriprive anemina are independent
H_a hypothesis	Sex and Feriprive anemina are dependent
alpha	0.05
df (degrees of freedom)	=(2-1)*(2-1), where 2 = number of rows and 2 = number of columns
Critical Chi-square	=CHISQ.INV.RT ((value of alpha),(value of degrees of freedom))
p-values	=CHISQ.TEST((the 4 values from observed table),(the 4 values from expected table))
Observed Chi-square	=CHISQ.INV.RT((select the cell where the p-value is),(select the cell where df are))

Statistical conclusion	
Clinical conclusion	

- *Alpha*: use alpha equal to 0.05 for this analysis (significance level of 5%).
- *df*: the formula for the degree of freedom (df) for the chi-square test of independence is df = (r - 1)·(c - 1), where r = the number of rows in the cross tabulation and c = the number of columns in the cross tabulation. Because our cross tabulation has 2 rows and 2 columns, df = (2 - 1)·(2 - 1) = 1.
- *Critical chi-square*: the CHIINV function will be use in order to find the critical chi-square value. Select the corresponding cell and **[Insert – Function... – Function category (Statistical) - CHIINV]**:
  - Click in the probability window of the CHIINV dialog box and enter the value of alpha
  - Deg_freedom: click in the Deg_freedom window and enter the df value by selecting the corresponding cell
- *p-value:* this is the probability associated with the observed value of chi-square. The CHITEST function will be use to find the p-value. Select the cell where you want the result and **[Insert – Function... – Function category (Statistical) – CHISQ.TEST]**. Fill the CHISQ.TEST dialog box with requested information:

Observed table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	13	48	61
M	16	33	49
Grand Total	**29**	**81**	**110**

Expected table	Feriprive anemia = yes	Feriprive anemia = no	Grand Total
F	16	45	61
M	13	36	49
Grand Total	**29**	**81**	**110**



Refers the range associated with the observed cell frequencies (just 4 cells)/

Refers the range associated to the expected frequencies (just 4 cells)/

- *Observed chi-square*: the CHIINV function will be use to find the observed chi-square value.
  - *Probability*: enter here p-value
  - *Deg_freedom*: enter the df value.
- The results will look like in the example bellow:

Chi-square test of independence: sex vs ferriprive anemia	
H_0 hypothesis	Sex and Feriprive anemina are independent
H_a hypothesis	Sex and Feriprive anemina are dependent
alpha	0.05
df (degrees of freedom)	1
Critical Chi-square	3.8415
p-values	0.1797
Observed Chi-square	1.8005
Statistical conclusion	Null hypothesis could not be rejected
Clinical conclusion	Sex and feriprive anemia are independent