

CHI SQUARE TEST

In order to do a Chi square test in Microsoft Excel, Libre Office Calc, Google Spreadsheet or any others similar software out there, the procedure is the same.

Step 1: Obtain the **Observed** frequencies table

In the present example I isolated the data from students.xlsx for a better view.

In order to obtain the observed frequencies from the below picture we use the **COUNTIFS** function.

E	F	G	H	I	
		Observed values table			
		Biostat difficult?			
		YES (1)	NO (2)	TOTAL	
	FEMALE (1)	6	15	21	
GENDER	MALE (2)	19	28	47	
	TOTAL	25	43	68	

		Biostat difficult?			
		YES (1)	NO (2)	TOTAL	
	FEMALE (1)	a	b	a+b	
GENDER	MALE (2)	c	d	c+d	
	TOTAL	a+c	b+d	a+b+c+d	

a represents the number of cases of **female** students who perceive the **biostatistics** to be **difficult**

b represents the number of cases of **female** students who perceive the **biostatistics not** to be **difficult**

c represents the number of cases of **male** students who perceive the **biostatistics** to be **difficult**

d represents the number of cases of **male** students who perceive the **biostatistics not** to be **difficult**

Column A contain the students` gender (1=female, 2=male)

Column B contain the students` perception about biostatistics difficulty (1=yes, 2=no)

a=countifs(A2:A69;1;B2:B69;1)

A2:A69 is the cells range for gender, and the criteria=1 (FEMALE);

B2:B69 is the cells range for biostatistics difficulty, and the criteria =1 (YES)

b=countifs(A2:A69;1;B2:B69;2)

A2:A69 is the cells range for gender, and the criteria=1 (FEMALE);

B2:B69 is the cells range for biostatistics difficulty, and the criteria =2 (NO)

c=countifs(A2:A69;2;B2:B69;1)

A2:A69 is the cells range for gender, and the criteria=2 (MALE);

B2:B69 is the cells range for biostatistics difficulty, and the criteria =1 (YES)

d=countifs(A2:A69;2;B2:B69;2)

A2:A69 is the cells range for gender, and the criteria=2 (MALE);

B2:B69 is the cells range for biostatistics difficulty, and the criteria =2 (NO)

To make the second step easier, we can calculate the totals of the rows and columns

Step 2: Obtain the Expected frequencies table

		Biostat difficult?	
		YES (1)	NO (2)
GENDER	FEMALE (1)	$(a+b)*(a+c)/(a+b+c+d)$	$(a+b)*(b+d)/(a+b+c+d)$
	MALE (2)	$(c+d)*(a+c)/(a+b+c+d)$	$(c+d)*(b+d)/(a+b+c+d)$

You can observe that for the female gender and the perception of biostatistics being difficult the expected frequency is obtained by multiplying the number of females by the number of biostatistics being difficult and then dived by the total number of our study sample.

To obtain the values for the rest of the cell we apply the same calculation, and finally we will obtain the results from the image below.

	K	L	M	N
		Expected values table		
			YES (1)	NO (2)
		FEMALE (1)	7,720588235	13,27941176
GENDER		MALE (2)	17,27941176	29,72058824

Step 3 Calculate the **p value** of the Chi-square test
in a empty cell use the CHITEST function, or enter the formula.

p value =CHITEST(G5:H6;M5:N6)

G5:H6 represents the values of observed frequencies table
M5:N6 represents the values of expected frequencies table

fx =CHITEST(G5:H6;M5:N6)									
	F	G	H	I	J	K	L	M	N
1		Observed values table					Expected values table		
2									
3		Biostat difficult?							
4		YES (1)	NO (2)	TOTAL			Q4		
5	FEMALE (1)	6	15	21			=CHITEST(G5:H6;M5:N6)	YES (1)	NO (2)
6	MALE (2)	19	28	47		GENDER	FEMALE (1)	7,720588235	13,27941176
7	TOTAL	25	43	68			MALE (2)	17,27941176	29,72058824
8									

The result is the same regardless of the chosen method (Excel, Spreadsheet or EpiInfo)

Chi square test
p value = 0,34893843

		Outcome		
		Yes	No	Total
Exposure	Yes	6	15	21
	Row %	28.57%	71.43%	100.00%
	Col %	24.00%	34.88%	30.88%
	Total	19	28	47
Exposure	No	19	28	47
	Row %	40.43%	59.57%	100.00%
	Col %	76.00%	65.12%	69.12%
	Total	25	43	68
Exposure	Total	25	43	68
	Row %	36.76%	63.24%	100.00%
	Col %	100.00%	100.00%	100.00%
	Total	25	43	68

Odds-based Parameters			
	Estimate	Lower	Upper
	0.5895	0.1939	1.7916
Ratio	0.5940	0.1821	1.7935
		0.1590	1.9964

Statistical Tests		
	X ²	2 Tailed P
Uncorrected	0.8773	0.34893843
Mantel-Haenszel	0.8644	0.35250690
Corrected	0.4415	0.50639506