



# Chi square - test, Odds Ratio

## Chi square - test, Odds Ratio

*Elias Zintzaras, M.Sc., Ph.D.*

*Professor in Biomathematics-Biometry  
Department of Biomathematics  
**School of Medicine**  
**University of Thessaly***

*Institute for Clinical Research and Health Policy Studies  
Tufts University School of Medicine  
Boston, MA, USA*

*Theodoros Mprotsis, MSc, PhD  
Teacher & Research Fellow  
**(<http://biomath.med.uth.gr>)**  
**University of Thessaly**  
**Email: [tmprotsis@uth.gr](mailto:tmprotsis@uth.gr)***



## Comparative analysis of drug A and drug B

Suppose in a study, 28 people were given either drug A or drug B. Of the 12 people who received drug A, 4 were cured, and of the 16 people who received drug B, 9 were cured. We want to test whether there is a difference between the two drugs and if there is a relationship between the drug and the therapeutic effect.

The data is presented in the form of a 2x2 matrix.

	Therapeutic effect	
Drug	Cured=1	Not cured=2
A=1	4	8
B=2	9	7



# Entering data and defining variables

Enter the data in the **Data View** and define the variables in the **Variable View**

	drug	outcome
1	1	1
2	1	1
3	1	1
4	1	1
5	1	2
6	1	2
7	1	2
8	1	2
9	1	2

10	1	2
11	1	2
12	1	2
13	2	1
14	2	1
15	2	1
16	2	1
17	2	1
18	2	1
19	2	1
20	2	1

21	2	1
22	2	2
23	2	2
24	2	2
25	2	2
26	2	2
27	2	2
28	2	2



# Running the chi-square test

- To test the relationship between the **two categorical variables** or compare the percentages of patients treated by the two drugs, we will apply the **chi-square test**.
- Choose **Analyze -> Descriptive Statistics -> Crosstabs** from the menu.

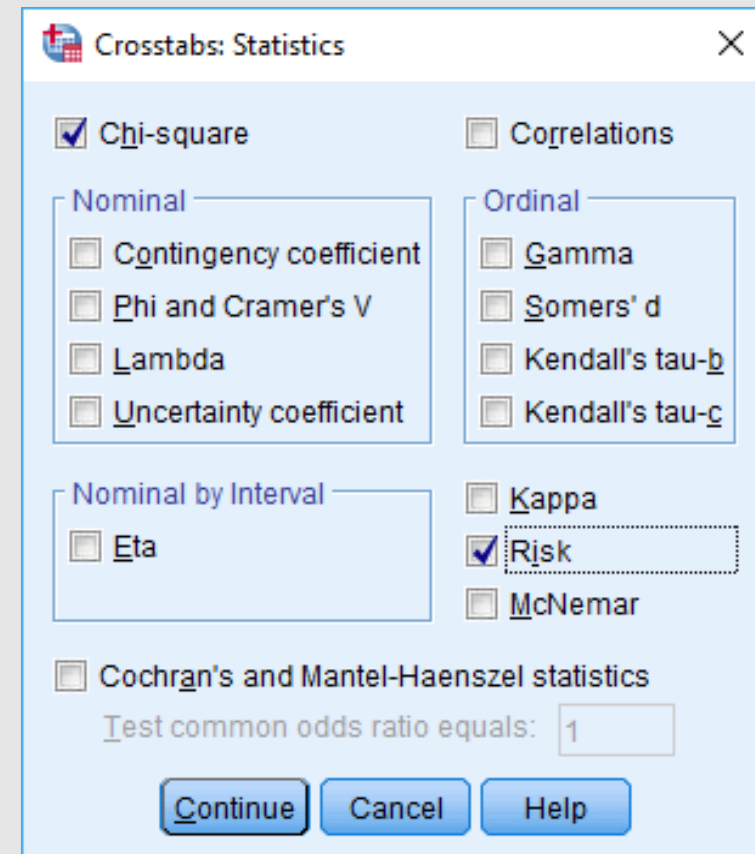
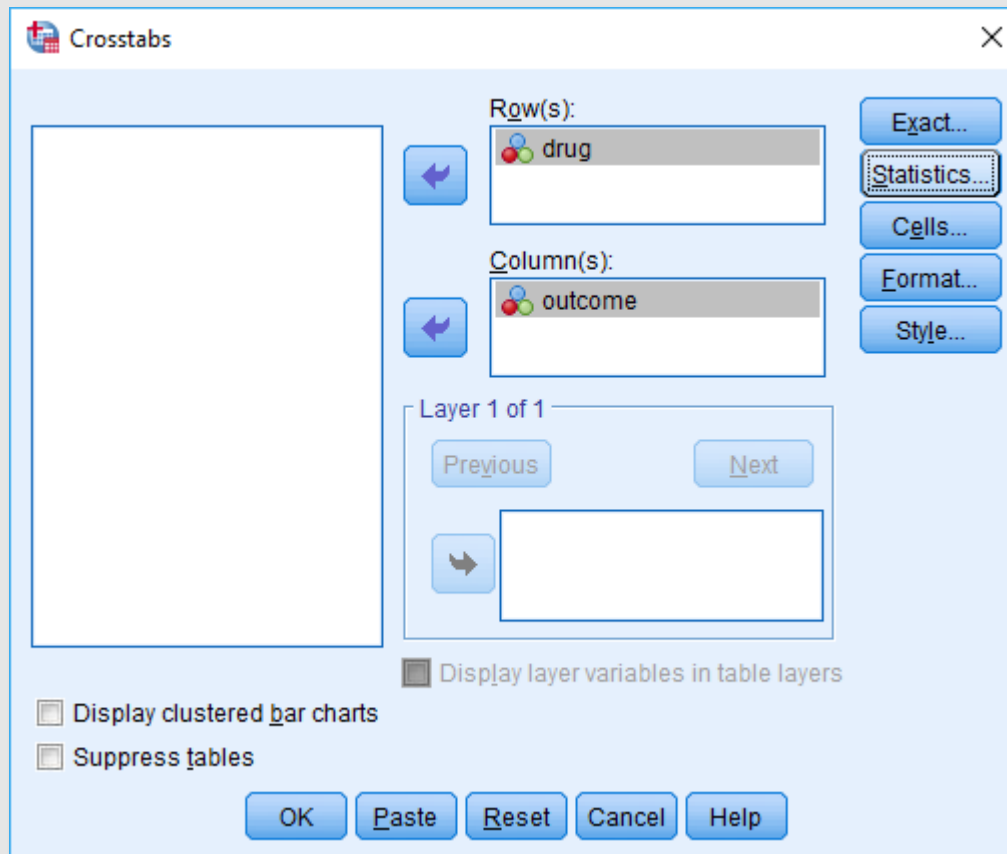
The screenshot shows the IBM SPSS Statistics Data Editor interface. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Extensions, and Windows. The 'Analyze' menu is open, showing options like Reports, Descriptive Statistics, Tables, Compare Means, General Linear Model, Generalized Linear Models, Mixed Models, Correlate, Regression, Loglinear, and Classify. The 'Descriptive Statistics' sub-menu is also open, showing options like Frequencies..., Descriptives..., Explore..., Crosstabs..., TURF Analysis, Ratio..., P-P Plots..., and Q-Q Plots... The 'Crosstabs...' option is highlighted. In the background, a data editor window is visible with columns for 'drug' and 'outcome' and rows of data.

	drug	outcome
1	1	1
2	1	1
3	1	1
4	1	1
5	1	2
6	1	2
7	1	2
8	1	2



# Running the chi-square test

Drag the variable **drug** from the left box into the **Row(s):** box, and the variable **outcome** from the left box into the **Column(s):** box. Then, choose **Statistics**, select **Chi-square** and **Risk**, and press **Continue**, and **OK**.





# Results and interpretation

In the **Chi-Square Tests** table, the **Pearson Chi-Square** value is **1.448**, which is not statistically significant ( $p > 0.05$ ), indicating that there is no relationship between the drug and treatment outcome. Alternatively, the percentages of patients treated by the two drugs do not differ.

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.448 <sup>a</sup>	1	.229		
Continuity Correction <sup>b</sup>	.673	1	.412		
Likelihood Ratio	1.467	1	.226		
Fisher's Exact Test				.276	.207
Linear-by-Linear Association	1.396	1	.237		
N of Valid Cases	28				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.57.

b. Computed only for a 2x2 table



# Results and interpretation

- In the **Risk Estimate table**, the **Odds Ratio** for drug (A/B) is 0.389.
  - This means that patients treated with drug A have 0.389 times the odds of being cured compared to patients treated with drug B.
  - It also means that drug A is associated with lower odds of treatment success compared to drug B.
  - In other words, patients on drug A are **less likely** to be cured than those on drug B. Specifically, the odds of being cured with drug A are about 61.1% lower than with drug B.
- The **95% confidence interval** for the **Odds Ratio** is 0.082 to 1.840. Since this interval **does include 1**, it suggests that the odds ratio is non statistically significant. The interval indicates that, with 95% confidence, the true odds ratio for drug A compared to drug B is between 0.082 and 1.840.
- For **easier interpretation**, we can calculate the **inverse** of the odds ratio ( $1/0.389 \approx 2.571$ ). This inverse value represents the odds of the comparison in the reverse direction (i.e., the odds of being cured with drug B compared to drug A). Thus, drug B has approximately 2.57 times the odds of success compared to drug A.
- This means that the 95% confidence interval for the odds ratio in the reverse direction ranges from approximately 0.543 ( $1/1.840 \approx 0.543$ ) to 12.20 ( $1/0.082 \approx 12.20$ ). Thus, drug B has between 0.543 and 12.20 times the odds of success compared to drug A.

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for drug (A / B)	.389	.082	1.840
For cohort outcome = Cured	.593	.239	1.471
For cohort outcome = Not Cured	1.524	.768	3.022
N of Valid Cases	28		

Chi-Square - test, OR with the use of weight cases





# Entering data and defining variables

Enter the data in the **Data View** and define the variables in the **Variable View** as shown below

	drug	outcome	freq	va
1	1	1	4.00	
2	1	2	8.00	
3	2	1	9.00	
4	2	2	7.00	
5				

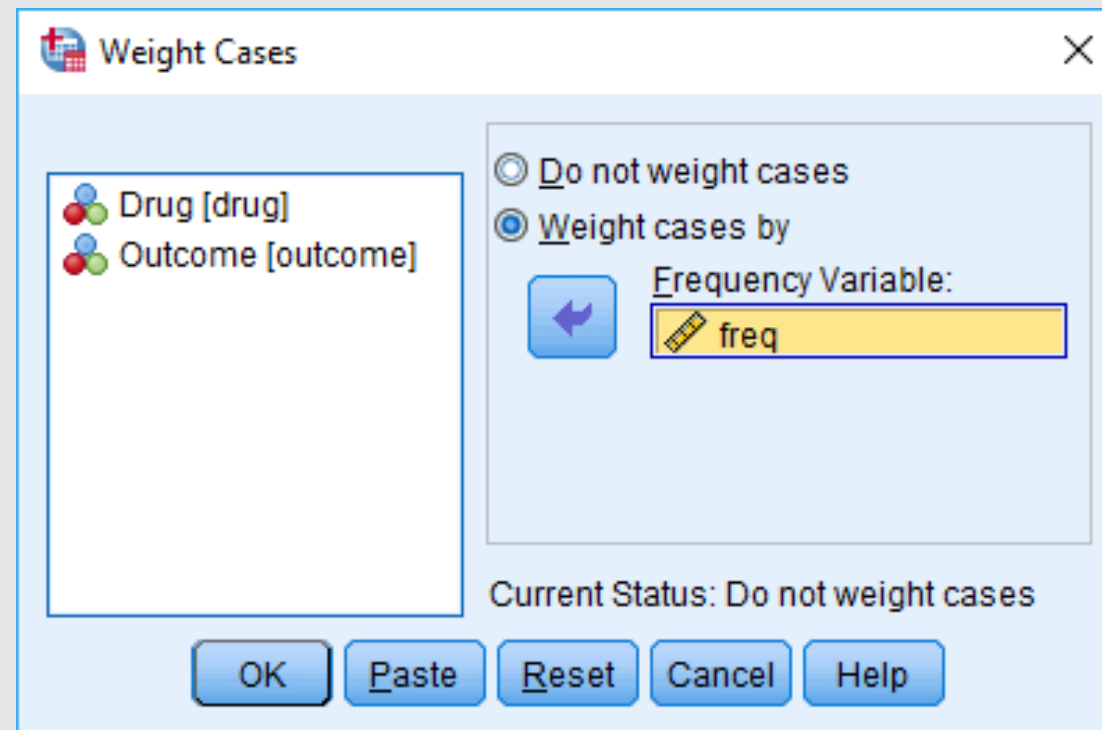
	Outcome	
Drug	Cured=1	Not cured=2
A=1	4	8
B=2	9	7

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
drug	Numeric	1	0	Drug	None	None	8	Right	Nominal
outcome	Numeric	1	0	Outcome	None	None	8	Right	Nominal
freq	Numeric	8	2		None	None	8	Right	Scale



# Weight Cases

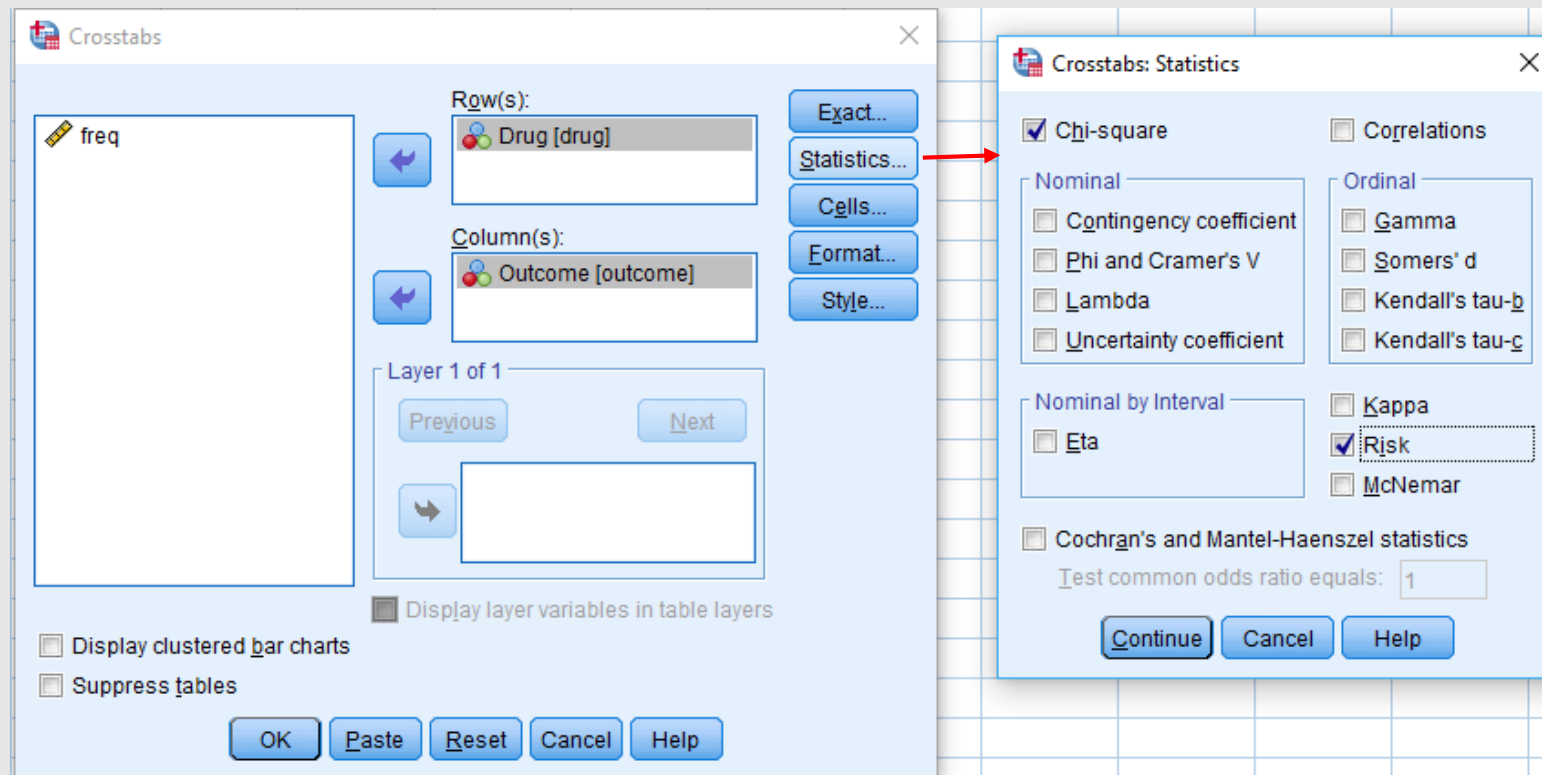
From the menu, select **Data -> Weight Cases**. In the window that appears, choose **Weight cases by**, and drag the variable **freq** into the **Frequency Variable:** field. Then, click **OK**.





# Chi-Square test, OR

- Choose **Analyze -> Descriptive Statistics -> Crosstabs** from the menu.
- Drag the variable **drug** from the left box into the **Row(s):** box, and the variable **outcome** from the left box into the **Column(s):** box. Then, choose **Statistics**, select **Chi-square** and **Risk**, and press **Continue**, and **OK**.





# Results

## Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.448 <sup>a</sup>	1	.229		
Continuity Correction <sup>b</sup>	.673	1	.412		
Likelihood Ratio	1.467	1	.226		
Fisher's Exact Test				.276	
Linear-by-Linear Association	1.396	1	.237		
N of Valid Cases	28				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.

b. Computed only for a 2x2 table

## Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for drug (A / B)	.389	.082	1.840
For cohort outcome = Cured	.593	.239	1.471
For cohort outcome = Not Cured	1.524	.768	3.022
N of Valid Cases	28		

Interactive Calculator for chi square tests of  
independence  
([https://biomath.med.uth.gr/statistics/chi\\_square.html](https://biomath.med.uth.gr/statistics/chi_square.html))

However, it is necessary to explicitly incorporate observed zero frequencies (you must enter "0" in such fields else it is presumed that they are not a part of your design). After entering your data, click Calculate and you should see results in the cells at the right. A pop up window will also merely indicates that p is very little.

The result is not significant at  $p < .05$ .  $X^2 (1, N = 28) = 1.448, p = 0.22884914$

	Col 1	Col 2	Col 3	Col 4	Col 5	Col 6	Col 7	Col 8	Col 9	Col 10	Total	Calculate	Reset all
Row 1	4	8									12		
Row 2	9	7									16		
Row 3													
Row 4													
Row 5													
Row 6													
Row 7													
Row 8													
Row 9													
Row 10													
Total	13	15									28		

Chi square:  
1.448

Degrees of freedom:  
1

P-value:  
0.22884914

Yates' chi-square:  
0.673

Yates' P-value:  
0.41200762

Edwards' chi-square:  
0.191

Edwards' P-value:  
0.66208563