DATAtab

CHI² TEST

Playbook

Theory & Example

Author: Dr. Mathias Jesussek

DATAtab

What is a **Chi^2 test?**

The **Chi^2 test** is a **hypothesis** test that is used when you want to determine. If there is a **relationship** between **two categorical variables**.

Categorical variables are, for example:

Gender

- Male
- Female

Preferred newspaper

- USA Today
- Wall Street Journal
- New York Times
- New York Post

Highest educational level

- Without graduation
- College
- Bachelor's degree
- Master's degree

No categorical variables are, for example:

Weight
of a person

Saraly
of a person

Power consumption

DATAtab

When do we use a Chi² test?

If we now have **two categorical variables** and we want to **test** whether there is a **relationship**, we use a **Chi² test**.

Example 1:

Is there a **relationship** between **gender** and the **preferred newspaper**?

Gender

1 = Male
2 = Female

Preferred newspaper

1 = USA Today
2 = Wall Street Journal
3 = New York Times
4 = New York Post

Example 2:

Is there a **relationship** between **preferred newspaper** and **highest educational level?**

Preferred newspaper

1 = USA Today
2 = Wall Street Journal
3 = New York Times
4 = New York Post

Highest educational level

1 = Without graduation
2 = College
3 = Bachelor's degree
4 = Master's degree

Keep in mind two things

1 The **assumption** for the **Chi² test** is that the **expected frequencies** per cell are greater than **5**.

2 The **Chi² test** uses only the **categories** and **not rankings**. However, in the case of the **highest educational level**, there is a **ranking** of categories.

If you want to account for **rankings**, check out our tutorials on Spearman correlation, Mann-Whitney U test or Kruskal-Wallis test.

Highest educational level

Without graduation
College
Bachelor's degree
Master's degree

Comparing two categorical variables is **only one use case** for the Chi2 test. There are also other use cases for this test.

How do we calculate a Chi^2 test?

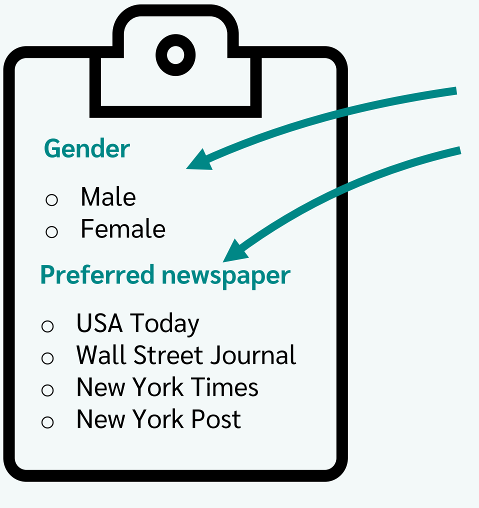
We would like to investigate whether **gender** has an influence on the **preferred newspaper**.

<u>So our question is:</u> Is there a **relationship** between **gender** and the **preferred newspaper?**

Null hypothesis:
There is **no relationship** between gender and the preferred newspaper.

Alternative hypothesis:
There **is a relationship** between gender and the preferred newspaper.

1 Create and send out a **questionnaire** that aks about **gender** und **preferred newspaper**.



2 The **results** of the survey are displayed in a **table**.

Case	Gender	Newspaper
1	Male	New York Post
2	Female	USA Today
3	Male	USA Today
4	Male	New York Times
5	Female	New York Times
6	Male	WSJ
7	Female	New York Post
...	...	...

In this table, we see **one respondent in each row.**

For example: The **sixth** respondent is **male** and states **Wall Street Journal** (WSJ)

3 We can now **copy** this table into a **statistics software** like **DATA*tab***, that then gives us the so-called **contingency table**.

Case	Gender	Newspaper
1	Male	New York Post
2	Female	USA Today
3	Male	USA Today
4	Male	New York Times
5	Female	New York Times
6	Male	WSJ
7	Female	New York Post
...	...	...

	Female	Male
USA Today	6	7
New York Post	13	16
New York Times	16	15
WSJ	8	11
Total	**43**	**49**

Variable **newspaper**

Variable **gender**

	Female	Male
USA Today	6	7
New York Post	13	16
New York Times	16	15
WSJ	8	11
Total	**43**	**49**

The **number** of times **each combination occurs** is plotted in the **cells**.

Example: In the survey there are **16 people** who stated **female** and New **York Times**.

4

Now we want to know if gender has an influence on the preferred newspaper. (= Is there a **relationship** between **gender** and the **preferred newspaper?**

CHI² TEST

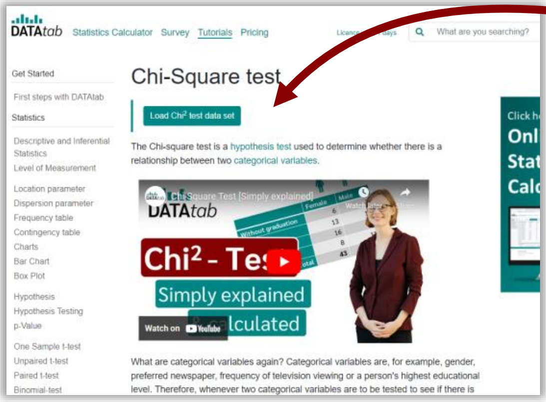
use a **statistical software** like **DATA***tab*

calculate **by hand** with the help of a **Chi² table**

Table of chi-squared distribution

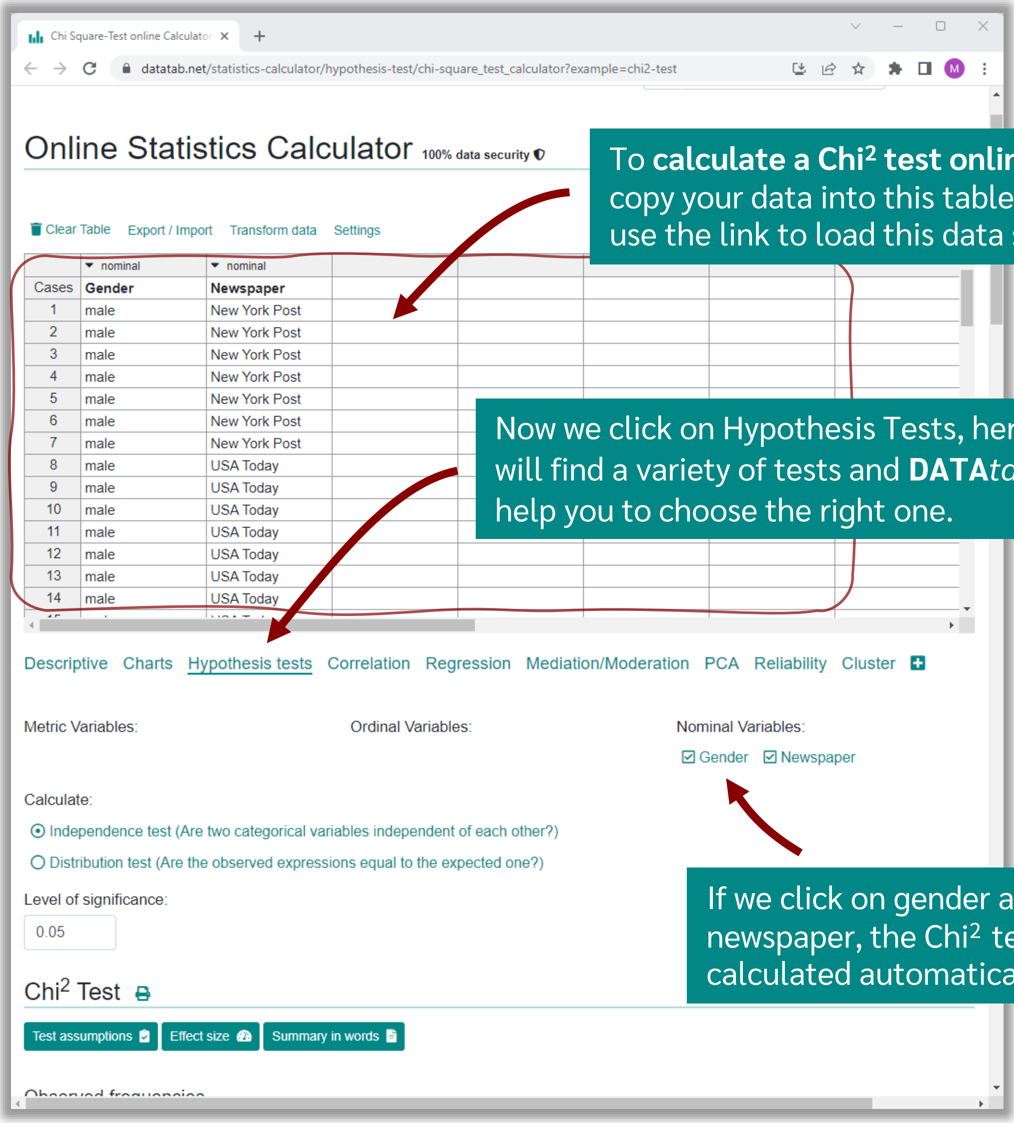
Significance level Alpha	0.995	0.975	0.2	0.1	0.05	0.025	0.02	0.01	0.005	0.002	0.001
Degrees of freedom											
1	0	0.001	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.55	10.828
2	0.01	0.051	3.219	4.605	5.991	7.378	7.824	9.21	10.597	12.429	13.816
3	0.072	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.86	16.924	18.467
5	0.412	0.831	7.289	9.236	11.07	12.833	13.388	15.086	16.75	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.69	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.18	11.03	13.362	15.507	17.535	18.168	20.09	21.955	24.352	26.124
9	1.735	2.7	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588

A Let's start with the uncomplicated variant and use **DATA*tab***!



If you like, you can **load** the sample **dataset** for calculation.

GO TO DATA*tab*



To **calculate a Chi² test online** copy your data into this table or use the link to load this data set.

Now we click on Hypothesis Tests, here you will find a variety of tests and **DATA*tab*** will help you to choose the right one.

If we click on gender and newspaper, the Chi² test will be calculated automatically.

Now we get the results for the Chi^2 test.

Chi^2 Test

Test assumptions | Effect size | Summary in words

We get the contingency table for the variables gender and newspaper. The contingency table shows us how often the respective combi-nations occur in our survey. Female and Wall Street Journal, for example, occurs 16 times.

Observed frequencies

Copy Word | Copy Excel

		Newspaper				
		New York Post	USA Today	New York Times	WSJ	Total
Gender	male	7	16	15	11	49
	female	6	13	16	8	43
	Total	13	29	31	19	92

Expected frequencies for perfectly independent variables

Copy Word | Copy Excel

		Newspaper				
		New York Post	USA Today	New York Times	WSJ	Total
Gender	male	6.92	15.45	16.51	10.12	49
	female	6.08	13.55	14.49	8.88	43
	Total	13	29	31	19	92

The second table shows how the contingency table should look like if the two variables were perfectly independent, i.e. if gender had no influence on the preferred newspaper.

For the assumptions to be fulfilled, it is important that All frequencies are greater than 5.

Chi-Square Test

Copy Word | Copy Excel

Chi^2	0.5
df	3
p	.918

The Chi^2 test now compares both table. The **p-value** is 0.91, which is much higher than our significance level of 0.05 and therefore we keep the null hypothesis.

If you don't know exactly how to interpret the results, just click on summary in words.

Chi^2 Test

Test assumptions | Effect size | Summary in words

Observed frequencies

Copy

Gen

Summary

A Chi^2 test was performed between Gender and Newspaper. All expected cell frequencies were greater than 5, thus the assumptions for the Chi^2 test were met. There was no statistically significant relationship between Gender and Newspaper, $\chi^2(3) = 0.5$, $p = .918$, Cramér's $V = 0.07$

The calculated p-value of .918 is above than the defined significance level of 5%. The Chi^2 test is therefore not significant and the null hypothesis is not rejected.

Expected frequencies for perfectly independent variables

Copy Word | Copy Excel

		Newspaper				
		New York Post	USA Today	New York Times	WSJ	Total
Gender	male	6.92	15.45	16.51	10.12	49
	female	6.08	13.55	14.49	8.88	43
	Total	13	29	31	19	92

Chi-Square Test

Copy Word | Copy Excel

Chi^2	0.5
df	3
p	.918

And now we come to the question of how to **calculate** the **Chi^2 test** by hand and we go through the **formulas** needed.

DATAtab

!

B How to **calculate** the **Chi² test** by hand: You need the contingency table with the observed frequencies and with the expected frequencies.

Observed frequencies

	Female	Male
USA Today	6	7
New York Post	13	16
New York Times	16	15
WSJ	8	11
Total	**43**	**49**

Expected frequencies

	Female	Male
USA Today	6.08	6.92
New York Post	13.55	15.45
New York Times	14.49	16.51
WSJ	8.88	10.12
Total	**43**	**49**

= frequencies, that would occur with perfectly independent variables

You can find how to calculate the expected frequencies on **DATA***tab* in the **tutorial** on the **Chi² test**.

We can now calculate the **Chi²** with this formula:

O_k **observed frequency**

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k}$$

The index k stands for the respective cell.

E_k **expected frequency**

Observed frequencies

	Female	Male
USA Today	6	7
New York Post	13	16
New York Times	16	15
WSJ	8	11
Total	**43**	**49**

Expected frequencies

	Female	Male
USA Today	6.08	6.92
New York Post	13.55	15.45
New York Times	14.49	16.51
WSJ	8.88	10.12
Total	**43**	**49**

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k} = \frac{(6-6.08)^2}{6.08} + \frac{(7-6.92)^2}{6.92} + \frac{(13-13.55)^2}{13.55} + \dots = \boxed{0.504}$$

plus the next cell

do this for all cells
and sum them up

Chi² value

This results in a
Chi² value of **0.504**

How do we calculate the **critical Chi^2 value?**

If the **calculated Chi^2 value** is larger than the **critical Chi^2 value**, there is a significant relationsship and the null hypothesis is rejected.

If we use a **statistical software**, we simply get a p-value displayed.

Chi-Square Test

Copy Word Copy Excel

Chi^2	0.5
df	3
p	.918

> 0.05

In our example case, the null hypothesis is **not rejected**.

! If the value is smaller than the **significance level**, e.g. 0.05, the **null hypothesis** is **rejected**, otherwise not.

By hand, however, you **can't** really **calculate** the **p-value**. Therefore, you read off in a **table** which **Chi^2 value** you would get with a **p-value** of **0.05**.

This **Chi^2 value** is called the **critical Chi^2 value**.

How do we calculate the **critical Chi² value?**

In order to calculate the **critical Chi² value**, we need the **degrees of freedom** (df).

These are obtained by taking the **number of rows minus 1**

times the **number of columns minus 1**

$$df = (\text{Number of rows - 1 }) \ (\text{Number of columns - 1 })$$

In our expamle
we have **4 rows** and **2 columns,**

	Female	**Male**
USA Today	6	7
New York Post	13	16
New York Times	16	15
WSJ	8	11
Total	**43**	**49**

therefore we get **3** times **1**

$$df = (4 - 1) \ (2 - 1) = \boxed{3}$$

and thus **3 degrees of freedom.**

!

Now let's take a look at the **table** of **critical Chi² values**

DATAtab

!

We have **3 degrees of freedom**

We select a **significance level of 0.05**

View table on datatab.net

Table of chi-squared distribution

Significance level Alpha	0.995	0.975	0.2	0.1	0.05	0.025	0.02	0.01	0.005	0.002	0.001
Degrees of freedom											
1	0	0.001	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.55	10.828
2	0.01	0.051	3.219	4.605	5.991	7.378	7.824	9.21	10.597	12.429	13.816
3	0.072	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.86	16.924	18.467
5	0.412	0.831	7.289	9.236	11.07	12.833	13.388	15.086	16.75	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.69	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.18	11.03	13.362	15.507	17.535	18.168	20.09	21.955	24.352	26.124
9	1.735	2.7	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588

Therefore, our **critical Chi² value is 7.815**.

The **critical Chi² value** of **7.815** is **larger** than our **calculated Chi² value.**

7.815 > 0.504

Thus, the **null hypothesis** is retained.

More about

CHI² TEST

on our website **datatab.net**

Feel free to share!