We investigate three groups of students first-year, second-year, and third year students. We want to find out whether the study results are dependent on the year. We have investigated 50 fist-year, 60 second year, and 40 third-year students and noted down the numbers of students with A, B, C, and D results in mathematics. The outcomes are summed up in the table below.

Y \ X	A	B	C	D	Σ
1st year	10	28	5	7	50
2nd year	18	30	6	6	60
3rd year	5	20	5	10	40
Σ	33	78	16	23	150

We will be testing the hypothesis H_0 that the random variables $X \in \{A, B, C, D\}$ and $Y \in \{1, 2, 3\}$ are independent at a 0.05 significance level. First we shall use the independency assumption to calculate the expected frequencies:

$Y \backslash X$	A	B	C	D	Σ
1st year	11	26	5.33	7.66	50
2nd year	13.2	31.2	6.4	9.2	60
3rd year	8.8	20.8	4.26	6.13	40
Σ	33	78	16	23	150

Here we use the formula

$$f_{ij} = \frac{y_i x_j}{T}$$

$Y \backslash X$	A	B	C	D	Σ
1st year	f_{11}	f_{12}	f_{13}	f_{14}	y_1
2nd year	f_{21}	f_{22}	f_{23}	f_{24}	y_2
3rd year	f_{31}	f_{32}	f_{33}	f_{34}	y_3
Σ	x_1	x_2	x_3	x_4	T

Now we can use the goodness-of-fit test and calculate the statistic t:

$$t = \sum_{x=1}^{4}\sum_{y=1}^{3}\frac{f_{xy}^{2}}{\overline{\overline{f}}_{xy}} - T = 7.489$$

To calculate the testing interval I_α we have to know the number of the degrees of freedom, which is the number of rows minus one times the number of columns minus one, that is, 6 in this particular example. Thus we have

$$I_\alpha = (0,12.592) \text{ and } t \in I_\alpha .$$

We reject H_0, which means that this test has not proved that the students' study results are dependent on their age

Two-by-two contingency tables

Sometimes we observe two population and want to determine whether the occurrence of a symptom depends on a particular population. This leads to two-by-two contingency tables, which are especially simple.

For example, we want to find out whether a new medicament is effective. We administer the medicament to a group of patients suffering from a disease that the new medicament is supposed to heal. We also administer a placebo to another group of patients and after a time measure the results. Suppose that we obtain the following results.

	Condition unchanged	Condition improved	Σ
Medicament	f_{11}	f_{12}	y_1
Placebo	f_{21}	f_{22}	y_2
Σ	x_1	x_2	T

If we hypothesize that the medicament is not effective, *X* and *Y* are independent and we get the following expected frequencies.

	Condition unchanged	Condition improved	Σ
Medicament	f_{11}	f_{12}	y_1
Placebo	f_{21}	f_{22}	y_2
Σ	x_1	x_2	T

To test this hypothesis we calculate the statistic

$$t = T\left(\frac{f_{11}^{2} x_2 y_2 + f_{12}^{2} x_1 y_2 + f_{21}^{2} x_2 y_1 + f_{22}^{2} x_1 y_1}{x_1 x_2 y_1 y_2} - 1 \right)$$

The chi-squared distribution has only one degree of freedom so we have

$$I_{0.01} = (0; 6.635) \quad \text{and} \quad I_{0.05} = (0; 3.841)$$

If $t \in I_{\alpha}$ we conclude that the test has not proved the medicament's effectiveness, otherwise it must be concluded that the test showed that the medicament has an influence over the patients' health condition.