Analyzing tables

Chi-squared, exact and monte carlo tests

Jon Michael Gran
Department of Biostatistics, UiO

MF9130 Introductory course in statistics
Thursday 26.05.2011

Overview

Aalen chapter 6.5, Kirkwood and Sterne chapter 17

- Chi-squared tests for 2x2 and larger contigency tables
- Exact test

Last lecture:

- Analysis of **proportions** - confidence intervals and tests
- The tests are not implemented in SPSS (but easily found other places)

Now:

- What if you have **more than two** exposure categories? Or outcomes?
- **Chi-square tests** are much more common and implemented in SPSS
- In case of little data: **exact tests**
- More on RR and OR

Chi-squared tests for 2×2 and larger contingency tables

$r \times c$ contigency tables

- The association between two categorical variables are displayed using $r \times c$ **contingency tables**, where r denotes the number of rows in the table and c the number of columns
- To examine whether there is an association between the row variable and the column variable, we use a **chi-squared test**

Chi-squared test for a 2×2 table

- The **test statistic** is

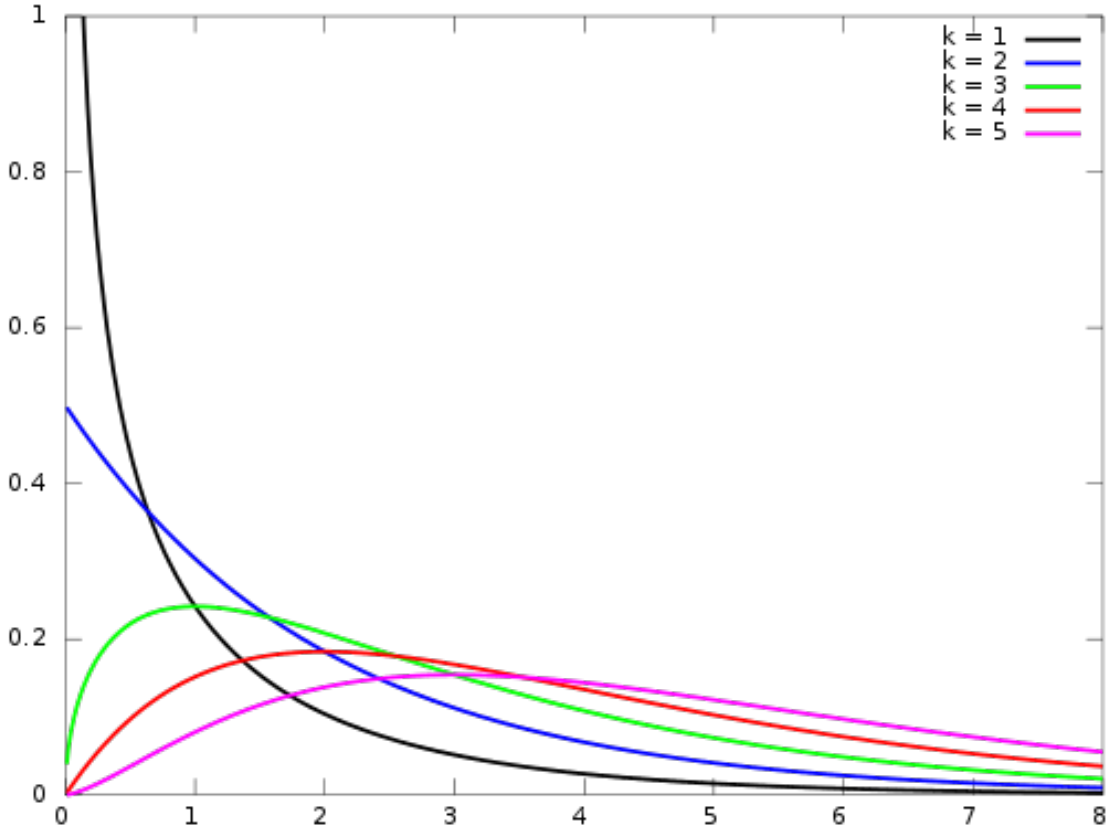
$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}, \quad \text{d.f.} = 1,$$

  where O_i and E_i denote the observed and expected values in the ith cell

- For a 2×2 table the test statistic is chi-squared distributed with 1 degree of freedom under the null hypothesis of no association between the two variables. This is equivalent to the z-**test** for the difference between two proportions

The logic behind the test...

- χ^2_n denotes a chi-square distribution with n degrees of freedom



- Identical to the sum of n squared normally distributed variables

Example 17.1 in Kirkwood & Sterne

We consider data from an **influenza vaccination trial**. In this case the exposure is vaccination (the row variable), so the table includes row percentages.

Observed numbers

	Influenza		
	Yes	**No**	**Total**
Vaccine	20 (8.3%)	220 (91.7%)	240
Placebo	80 (36.4%)	140 (63.6%)	220
Total	100 (21.7%)	360 (78.3%)	460

We want to assess the strength of the evidence that vaccination affected the probability of getting influenza.

We start by calculating the **expected numbers** under the assumption of no association between vaccination and subsequent contraction of influenza.
Overall 100/460 people got influenza, so the expected numbers **getting influenza** are:

- $100/460 \times 240 = 52.2$ in the vaccine group, and
- $100/460 \times 220 = 47.8$ in the placebo group.

Further, overall 360/460 people escaped influenza, so the expected numbers **escaping influenza** are:

- $360/460 \times 240 = 187.8$ in the vaccine group, and
- $360/460 \times 220 = 172.2$ in the placebo group.

The **test statistic** is

$$\begin{aligned}\chi^2 &= \frac{(20-52.2)^2}{52.2} + \frac{(80-47.8)^2}{47.8} + \frac{(220-187.8)^2}{187.8} \\ &\quad + \frac{(140-172.2)^2}{172.2} \\ &= 19.86 + 21.69 + 5.52 + 6.02 = 53.09,\end{aligned}$$

and the corresponding P-**value** is < 0.001. There is strong evidence against the null hypothesis of no effect of the vaccine on the probability of contracting influenza. It is therefore concluded that the vaccine is effective.

Alternative formulation of the Chi-squared test for a 2×2 table

- A quicker formula for calculating the **test statistic** on a 2×2 table is

$$\chi^2 = \frac{n \times (d_1 \times h_0 - d_0 \times h_1)^2}{d \times h \times n_1 \times n_0}, \quad \text{d.f.} = 1,$$

  using the standard notation for a 2×2 table

Example 17.1 in Kirkwood & Sterne

In the example of the **influenza vaccination trial**, the chi-squared is

$$\chi^2 = \frac{460 \times (20 \times 140 - 80 \times 220)^2}{100 \times 360 \times 240 \times 220} = 53.01,$$

which, apart from rounding error, is the same as the value obtained using the formula of observed and expected numbers

Test validity

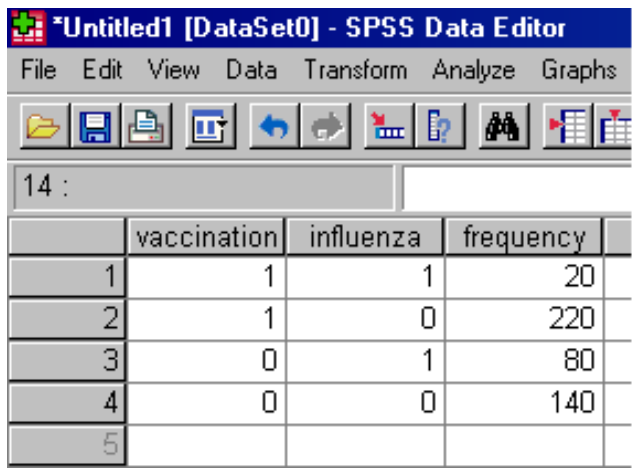
- The **chi-squared test** is valid when:
  - The overall total is more than 40, regardless of the expected values, or
  - The overall total is between 20 and 40 provided all the expected values are at least 5
- The use of the **exact test** is recommended when:
  - The overall total of the table is less than 20, or
  - The overall total is between 20 and 40 and the smallest of the four expected numbers is less than 5

Important

- The chi-squared test produce **only one p-value**
- Often a meassure of the effect with confidence intervals are required when publishing (for instance odds ratio, relative risk or the risk difference)

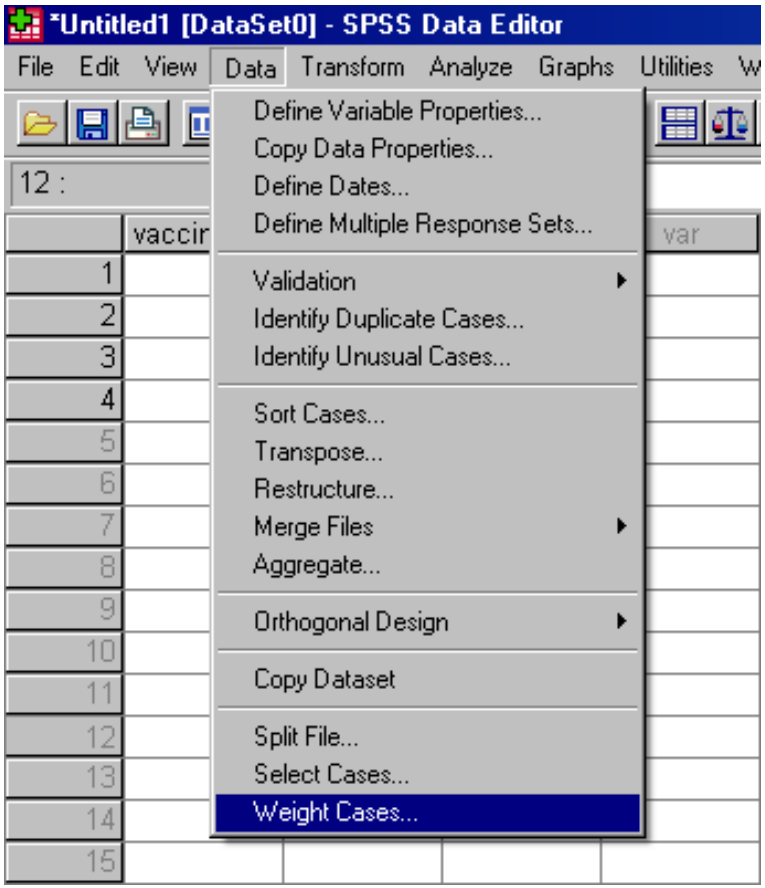
SPSS: Example 17.1 in Kirkwood & Sterne

1. Define the variables needed for the analysis

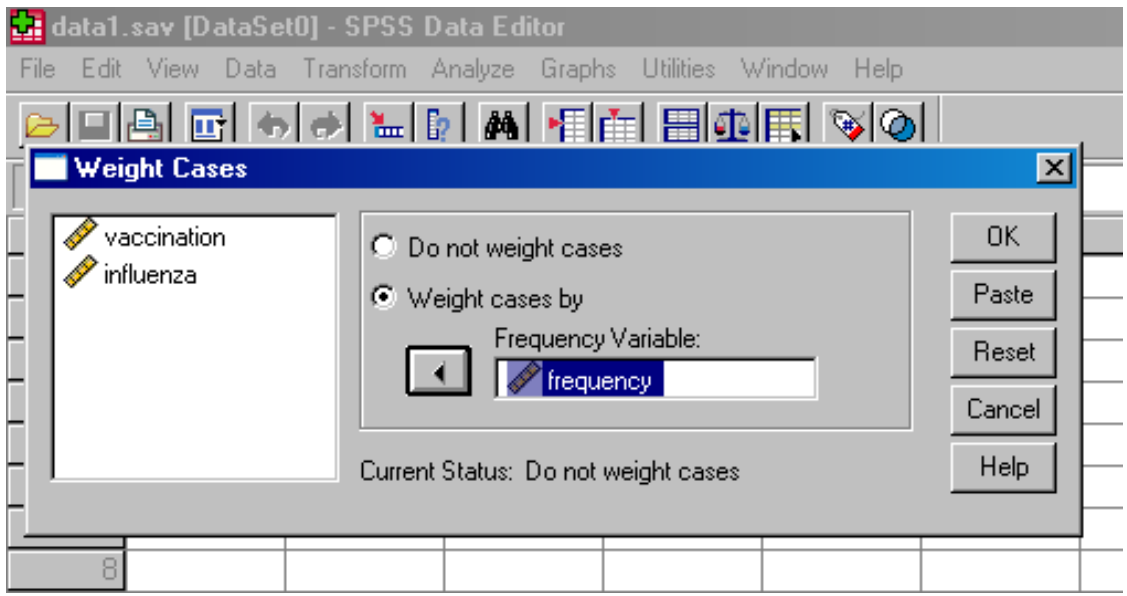


	vaccination	influenza	frequency
1	1	1	20
2	1	0	220
3	0	1	80
4	0	0	140
5			

2. Click on ***Data*** $\rightarrow$ ***Weight Cases...*** on the menu bar

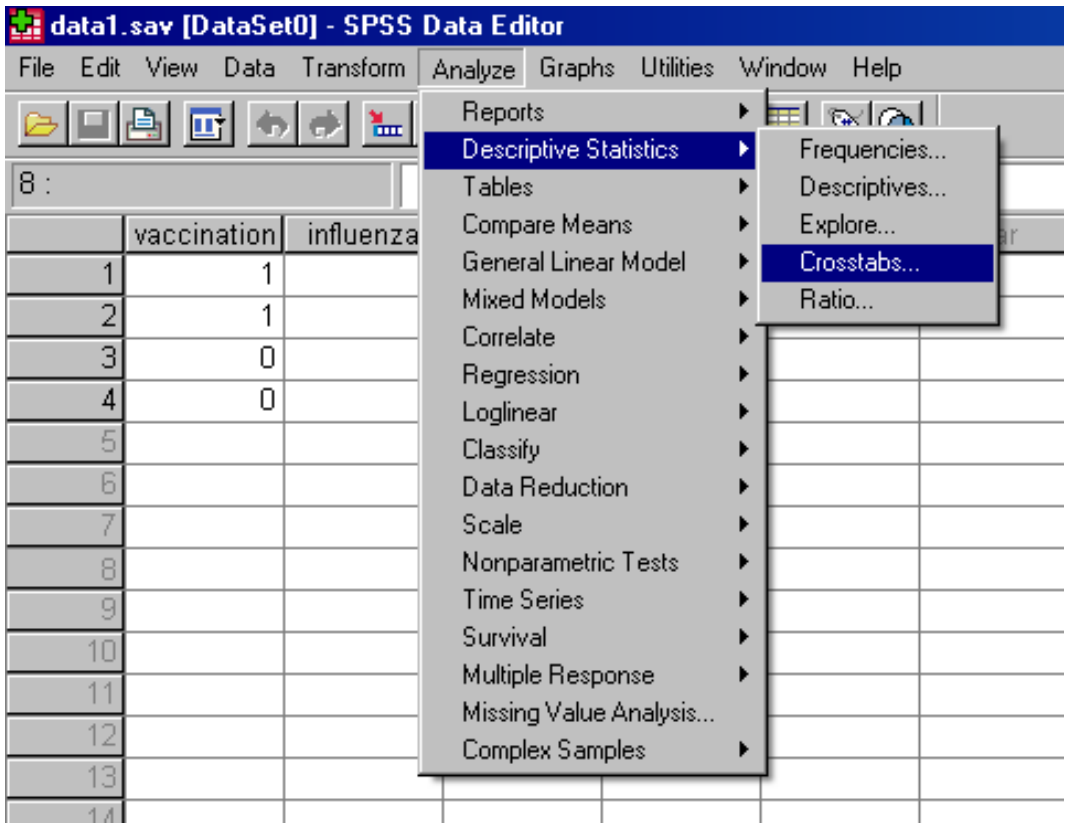


3. Check off ***Weight cases by*** and move the frequency variable (`frequency`) over to the empty field by marking it and clicking on the arrow

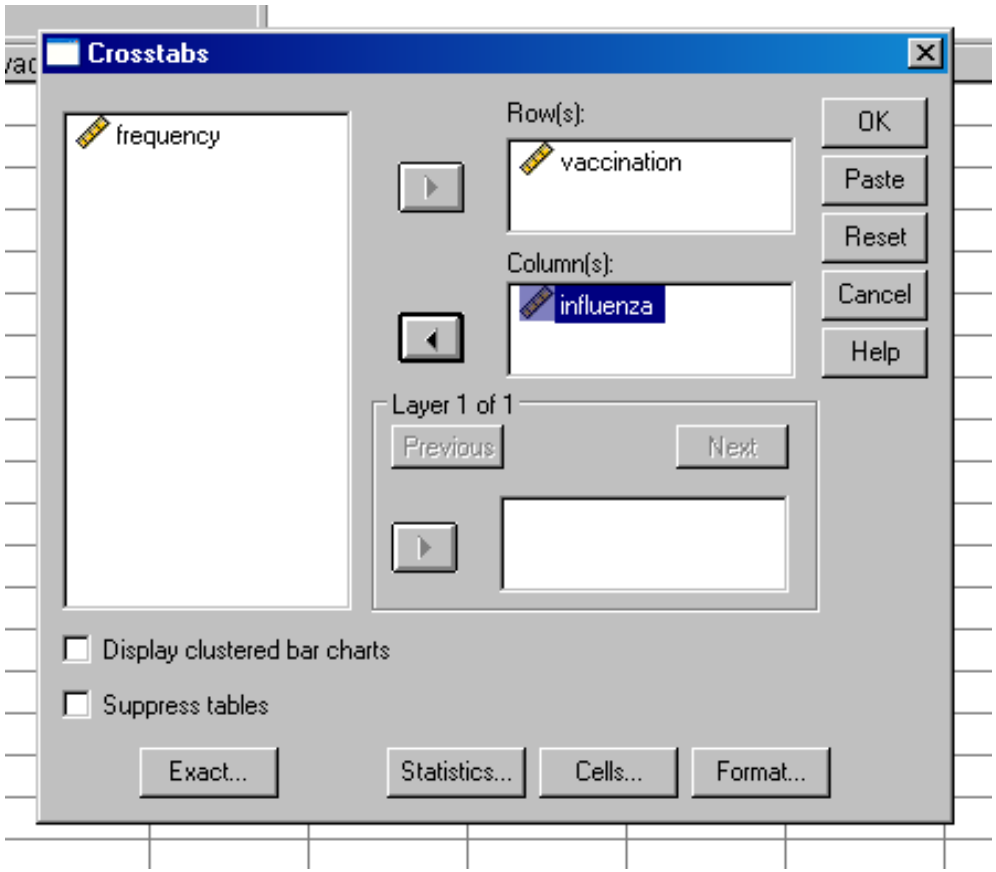


4. Click on ***OK***

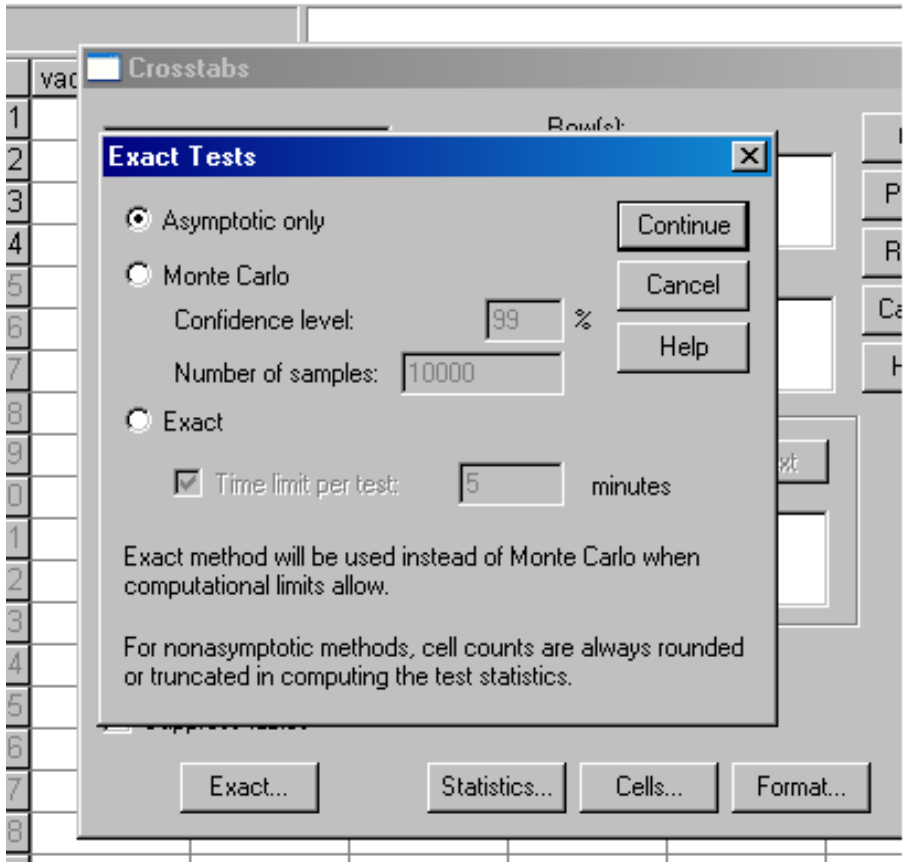
5. Click on ***Analyze*** → ***Descriptive Statistics*** → ***Crosstabs...*** on the menu bar



6. Move the exposure variable (`vaccination`) and the outcome variable (`influenza`) over to the empty fields by marking them and clicking on the arrows



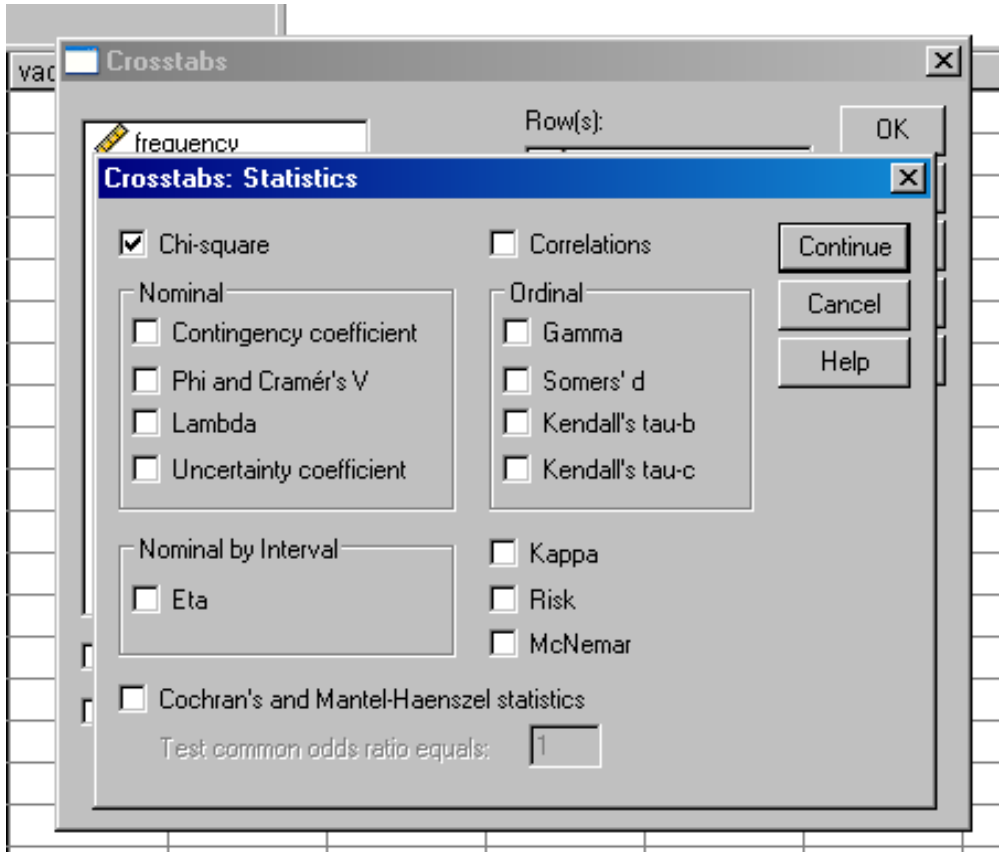
7. Click on ***Exact...***.



8. Check off ***Asymptotic only*** and click on ***Continue***

SPSS, cont.

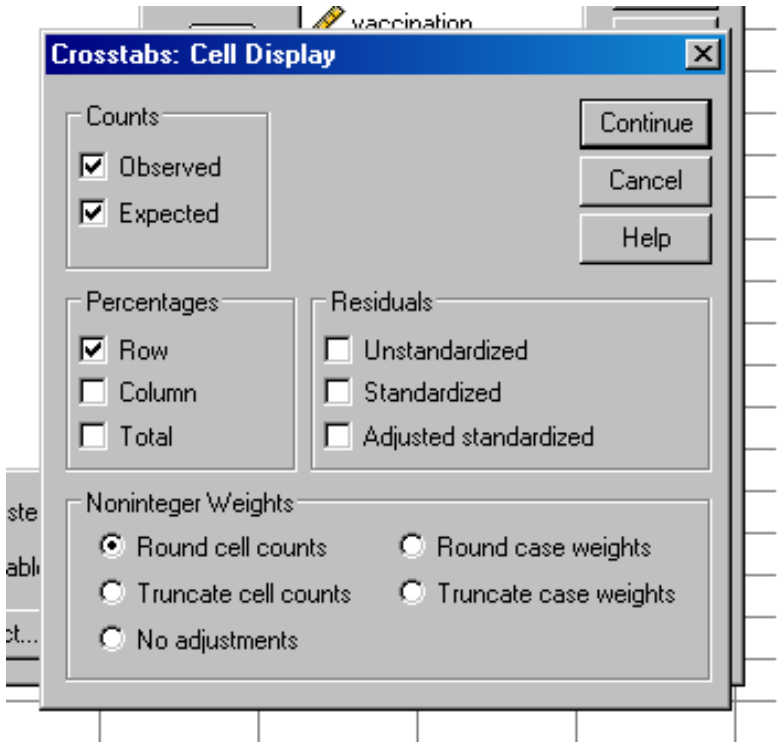
9. Click on ***Statistics...***



10. Check off ***Chi-square*** and click on ***Continue***

SPSS, cont.

11. Click on ***Cells...***



12. Check off both ***Observed*** and ***Expected*** under ***Counts*** and ***Row*** under ***Percentages*** and click on ***Continue***

13. Click on ***OK***
14. Interpret the results

vaccination * influenza Crosstabulation

			influenza		
			0	1	Total
vaccination	0	Count	140	80	220
		Expected Count	172,2	47,8	220,0
		% within vaccination	63,6%	36,4%	100,0%
	1	Count	220	20	240
		Expected Count	187,8	52,2	240,0
		% within vaccination	91,7%	8,3%	100,0%
Total		Count	360	100	460
		Expected Count	360,0	100,0	460,0
		% within vaccination	78,3%	21,7%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	53,008[b]	1	,000		
Continuity Correction [a]	51,374	1	,000		
Likelihood Ratio	55,606	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	52,893	1	,000		
N of Valid Cases	460				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 47,83.

(Fisher's) Exact test for 2×2 tables

- When the numbers in the 2×2 table are very small, we need an **exact test** to compare two proportions
- This is based on calculating the **exact probabilities** of the observed table and of more extreme tables with the same row and column totals, using the following formula:

$$\text{Exact probability} = \frac{d! \times h! \times n_1! \times n_0!}{n! \times d_1! \times d_0! \times h_1! \times h_0!}, \qquad (1)$$

with the standard notation for a 2×2 table

P-values in exact test for 2×2 tables

- There are *two* approaches to calculate the *P*-value:
  - *P*-value (approach I) = probability of observed table +probability of less probable tables
  - *P*-value (approach II) = $2 \times$ (probability of observed table +probability of more extreme tables in the same direction)

Example: 17.2 in Kirkwood & Sterne

Consider the results from a study to compare two treatment regimes for **controlling bleeding** in haemophiliacs ('*bløder*e') undergoing surgery

	Bleeding complications		
Treatment regime	**Yes**	**No**	**Total**
A (group 1)	1 (d_1)	12 (h_1)	13 (n_1)
B (group 0)	3 (d_0)	9 (h_0)	12 (n_0)
Total	4 (d)	21 (h)	25 (n)

Only one (8%) of the 13 haemophiliacs given treatment regime A suffered bleeding complications, compared to three (25%) of the 12 given regime B

These numbers are too small for the **chi-squared test** to be valid:

- the overall total, 25, is less than 40, and
- the smallest expected value, $4/25 \times 12 = 1.9$ (complications with regime B), is less than 5

The **exact test** should therefore be used

The **exact probability** of the observed table is

$$\begin{aligned} \text{Exact probability} &= \frac{4! \times 21! \times 13! \times 12!}{25! \times 1! \times 3! \times 12! \times 9!} \\ &= 0.2261 \end{aligned}$$

In addition, we need to calculate the probability that a **more extreme table** (with the same row and column totals as the observed table) could occur by chance under the null hypothesis that there is no difference between the two treatment regimes

			Total
	0	13	13
	4	8	12
Total	4	21	25

P=0.0391

			Total
	1	12	13
	3	9	12
Total	4	21	25

P=0.2261

			Total
	2	11	13
	2	10	12
Total	4	21	25

P=0.4070

			Total
	3	10	13
	1	11	12
Total	4	21	25

P=0.2713

			Total
	4	9	13
	0	12	12
Total	4	21	25

P=0.0565

According to **approach** I the total probability needed for the *P*-**value** is:

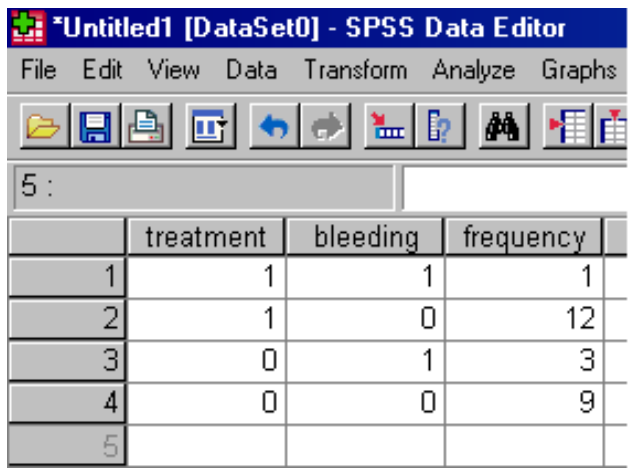
$$0.2261 + 0.0391 + 0.0565 = 0.3217,$$

and so there is clearly no evidence against the null hypothesis of no difference between the regimes. According to **approach** II the *P*-**value** obtained is:

$$2 \times (0.0391 + 0.2261) = 0.5304.$$

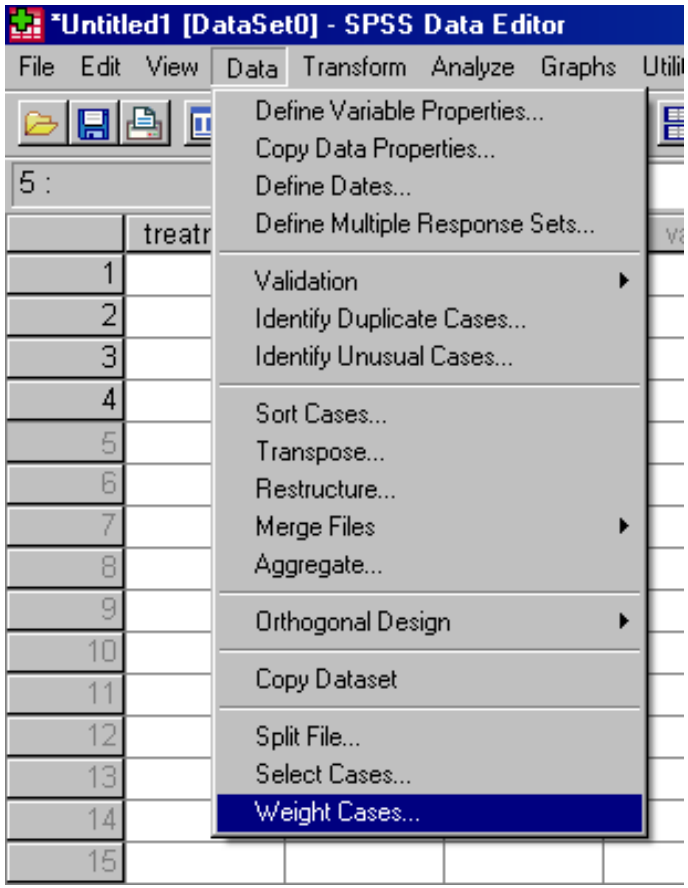
SPSS: Example 17.2 in Kirkwood & Sterne

1. Define the variables needed for the analysis

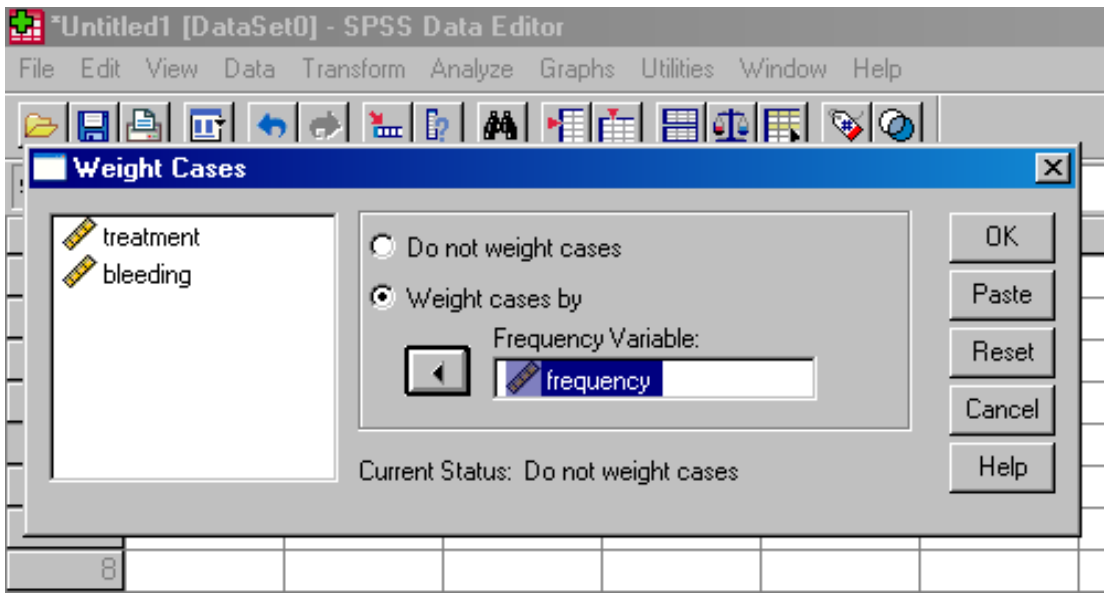


	treatment	bleeding	frequency
1	1	1	1
2	1	0	12
3	0	1	3
4	0	0	9
5			

2. Click on ***Data*** $\rightarrow$ ***Weight Cases...*** on the menu bar

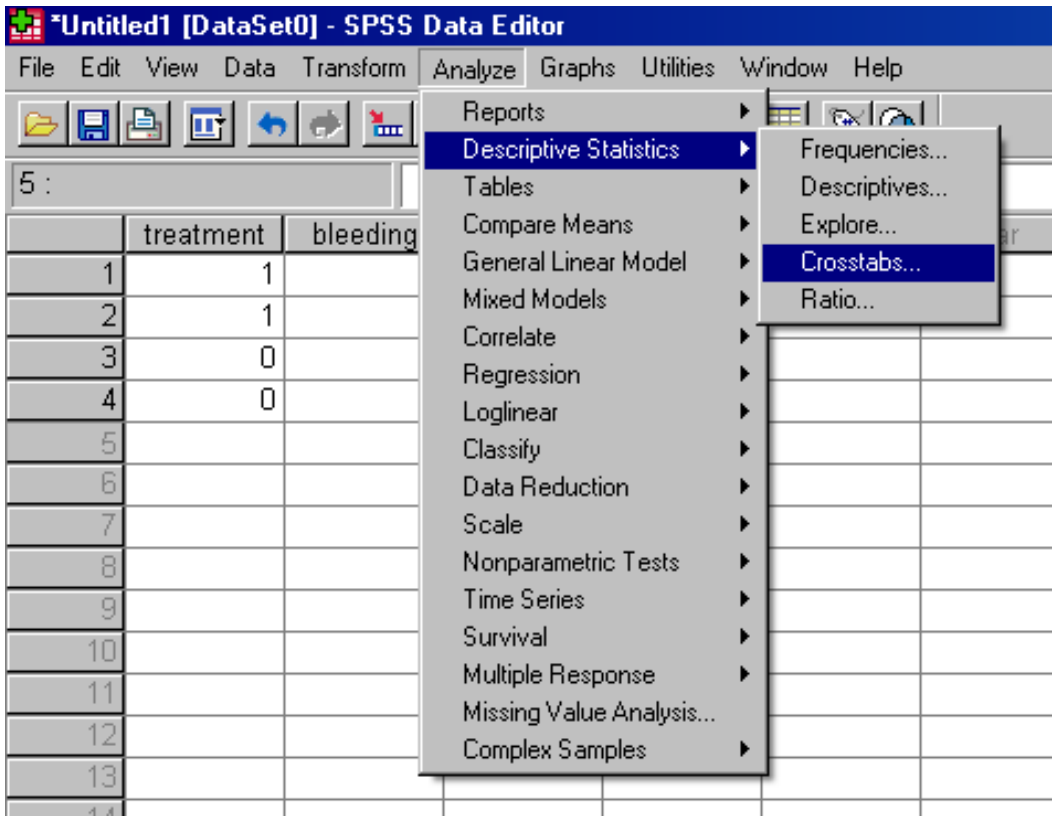


3. Check off ***Weight cases by*** and move the frequency variable (`frequency`) over to the empty field by marking it and clicking on the arrow

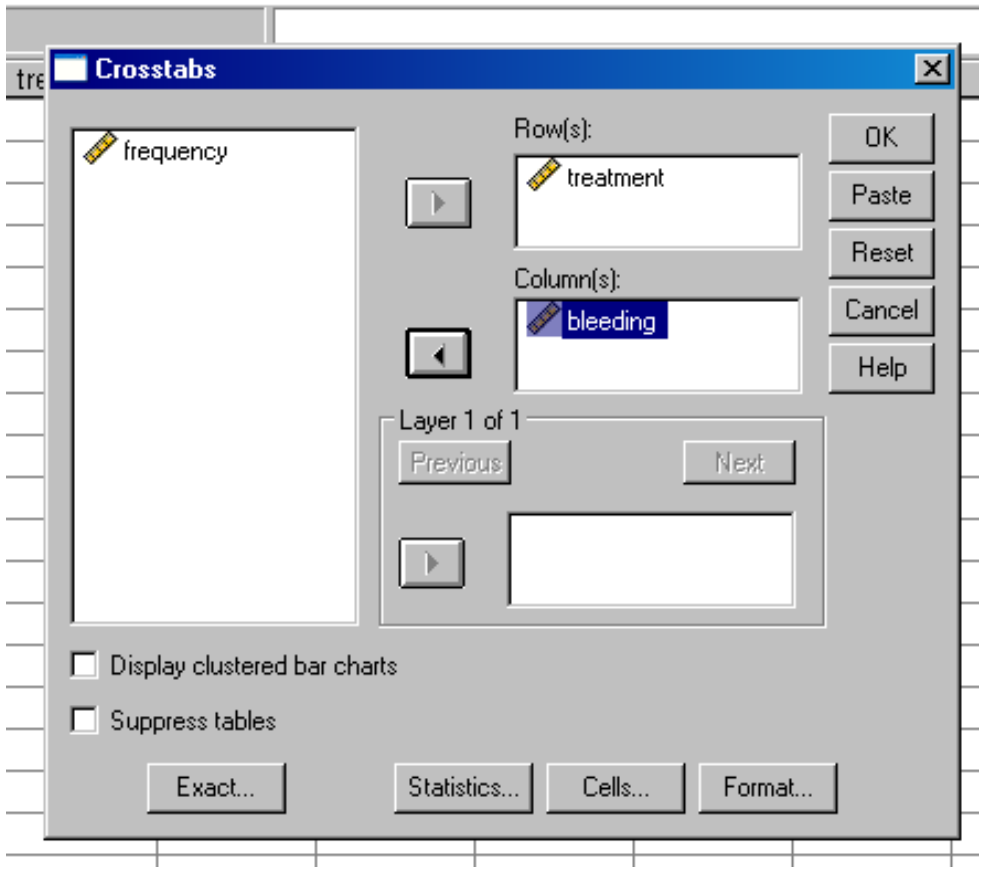


4. Click on ***OK***

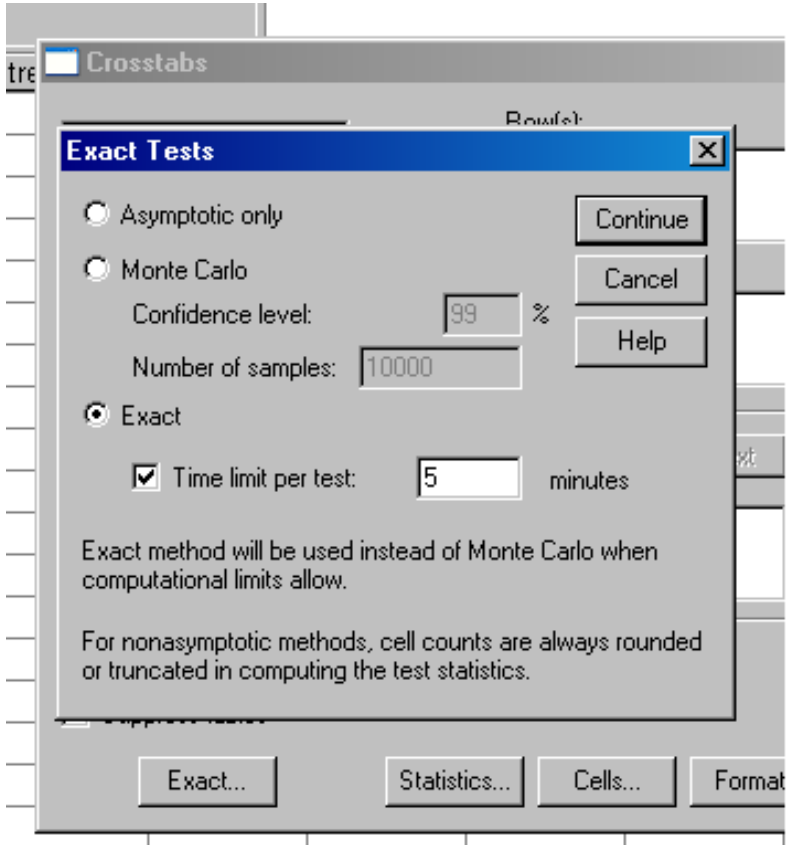
5. Click on ***Analyze*** → ***Descriptive Statistics*** → ***Crosstabs...*** on the menu bar



6. Move the exposure variable (`treatment`) and the outcome variable (`bleeding`) over to the empty fields by marking them and clicking on the arrows

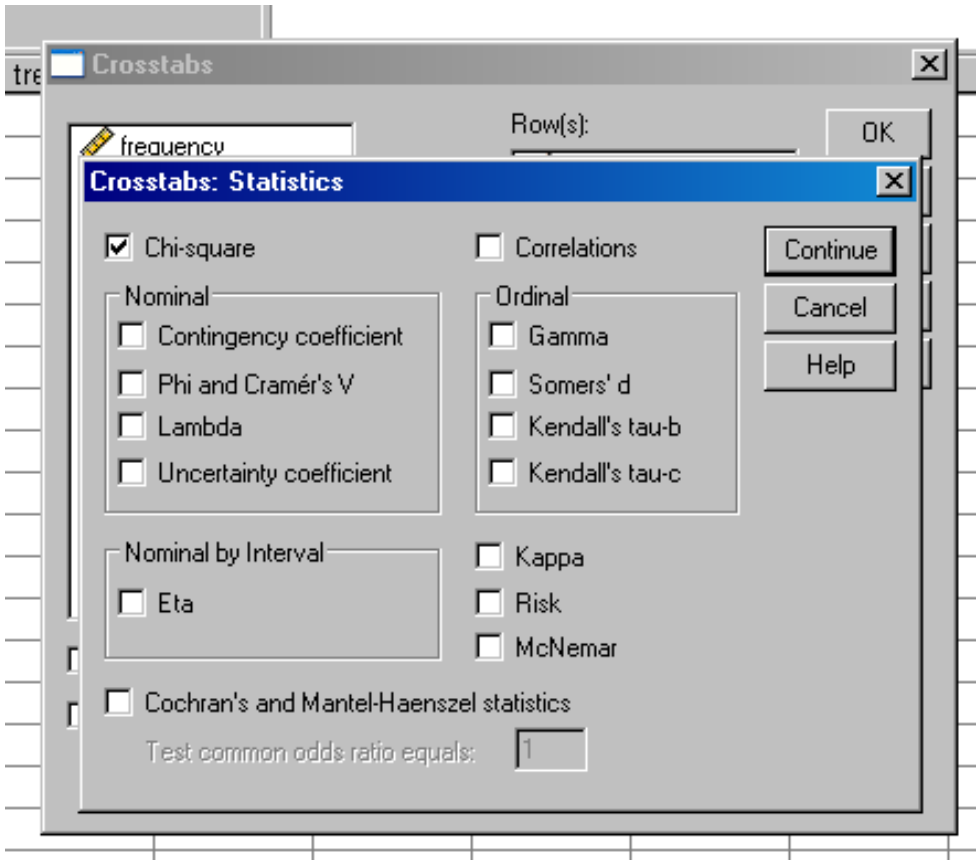


7. Click on ***Exact...***.



8. Check off ***Exact*** and click on ***Continue***

9. Click on ***Statistics...***



10. Check off ***Chi-square*** and click on ***Continue***

11. Click on ***Cells...***

12. Check off ***Observed*** under ***Counts*** and ***Row*** under ***Percentages*** and click on ***Continue***

13. Click on ***OK***
14. Interpret the results.

treatment * bleeding Crosstabulation

			bleeding		
			0	1	Total
treatment	0	Count	9	3	12
		% within treatment	75,0%	25,0%	100,0%
	1	Count	12	1	13
		% within treatment	92,3%	7,7%	100,0%
Total		Count	21	4	25
		% within treatment	84,0%	16,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1,391[b]	1	,238	,322	,265	
Continuity Correction [a]	,401	1	,527			
Likelihood Ratio	1,437	1	,231	,322	,265	
Fisher's Exact Test				,322	,265	
Linear-by-Linear Association	1,335[c]	1	,248	,322	,265	,226
N of Valid Cases	25					

a. Computed only for a 2x2 table

b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 1,92.

c. The standardized statistic is -1,155.

Larger contingency tables

- The **chi-squared test** can also be applied to larger tables, generally called $r \times c$ **tables**, where r denotes the number of rows in the table and c the number of columns
- The **test statistic** is:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}, \quad \text{d.f.} = (r-1) \times (c-1), \tag{2}$$

  which is chi-squared distributed with $(r-1) \times (c-1)$ degrees of freedom under the null hypothesis
- The general rule for calculating an **expected number** is:

$$E = \frac{\text{column total} \times \text{row total}}{\text{overall total}} \tag{3}$$

Test validity

- The approximation of **the chi-squared test is valid when**:
  - less than 20% of the expected numbers are under 5, and
  - none of the expected numbers is less than 1
- Sometimes this restriction **can be overcome by combining rows (or columns)** with low expected numbers, providing that these combinations make biological sense

Example: 17.3 in Kirkwood & Sterne

Consider the results from a survey to compare the **principal water sources** in three villages in West Africa.
The numbers of households using a river, a pond, or a spring are given. We will treat the **water source** as outcome and **village** as exposure, so column precentages are displayed.

Observed numbers

	Water source			
Village	**River**	**Pond**	**Spring**	**Total**
A	20 (40.0%)	18 (36.0%)	12 (24.0%)	50 (100.0%)
B	32 (53.3%)	20 (33.3%)	8 (13.3%)	60 (100.0%)
C	18 (45.0%)	12 (30.0%)	10 (25.0%)	40 (100.0%)
Total	70 (46.7%)	50 (33.3%)	30 (20.0%)	150 (100.0%)

Overall, 70 of the 150 households use a river. If there were no difference between villages, one would expect this same proportion of river usage in each village. Thus the **expected numbers** of households using a river in villages A, B and C, respectively, are:

$$\frac{70}{150} \times 50 = 23.3, \quad \frac{70}{150} \times 60 = 28.0 \quad \text{and} \quad \frac{70}{150} \times 40 = 18.7.$$

We use the same procedure to calculate the expected numbers of households using a pond and a spring in the villages.

Expected numbers

	Water source			
Village	**River**	**Pond**	**Spring**	**Total**
A	23.3	16.7	10.0	50
B	28.0	20.0	12.0	60
C	18.7	13.3	8.0	40
Total	70	50	30	150

The observed value of the **test statistic** is:

$$\begin{aligned}\chi^2 &= \frac{(20-23.3)^2}{23.3} + \frac{(18-16.7)^2}{16.7} + \frac{(12-10.0)^2}{10.0} \\ &\quad + \frac{(32-28.0)^2}{28.0} + \frac{(18-18.7)^2}{18.7} + \frac{(20-20.0)^2}{20.0} \\ &\quad + \frac{(8-12.0)^2}{12.0} + \frac{(12-13.3)^2}{13.3} + \frac{(10-8.0)^2}{8.0} \\ &= 3.53,\end{aligned}$$

with $\mathrm{d.f.} = (r-1) \times (c-1) = 2 \times 2 = 4$ degrees of freedom.

The corresponding ***P*-value** is 0.47. This means that there is no evidence of a difference between the villages in the proportion of households using different water sources.

More on OR and RR in SPSS

- RR can only be found for 2x2 tables in SPSS
- OR can be found generally, but have to use logistic regression (not a part of this curriculum)

McNemars test

- A special case of the 2x2 table for **paired categorical data**
- For example measuring the presence/absence of soemthing at to time points for each individual
- Check McNemar instead of Chi-squared in crosstabs in SPSS

Monte carlo tests

- Something between chi-squared tests and exact tests
- SPSS has the option Monte Carlo
- Based on simulations and less computationally intensive than the exact test
- Hardly relevant anymore, since Fisher's exact test can be used

Summary

Key words

- Contigency tables
- Chi-squared tests for 2×2 and $r \times c$ tables
- Exact tests

Notation

- χ^2_n
- O_i and E_i
- n, d, h