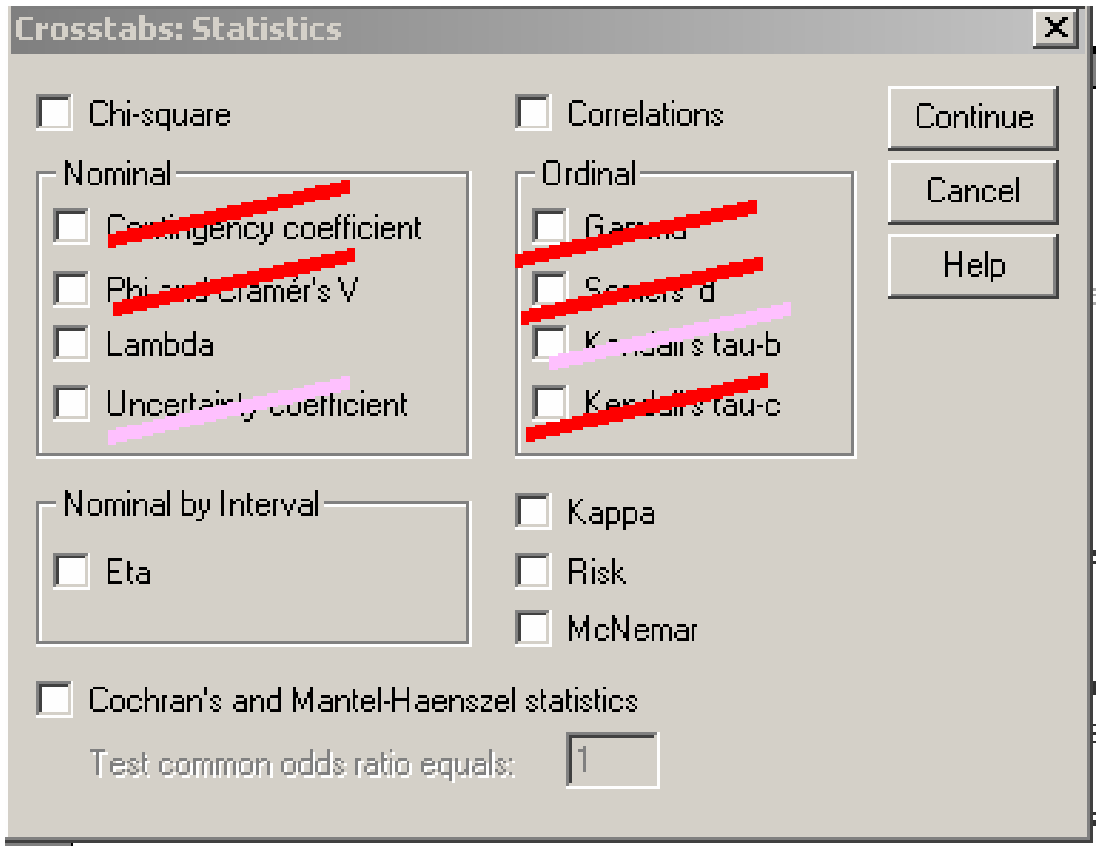
Categorical Analysis Contents

Which tests to use: Picture?



Recommended: Overview

Fisher's Exact test Pearson Chi-Square χ^2 Likelihood ratio		Tests for differences in proportions. Exact test (SPSS: 2 x 2 ; SAS: r x c) Tests for differences in proportions. Asymptotic test. H_0: Independency. ($\leq$ 20% cells expected freq. < 5) f(observed/expected) via χ^2		
Spearman correlation ρ_s	[-1;1]	Monotone correlation. Invariant against any monotone transformation. Two ordinal variables.		
κ (Cohen's Kappa)	[0;1]	Measure of concordance or reliability between n observers (treatments) for nominal or ordinal data (nxn)		
λ and Goodman-Kruskal tau	[0;1]	Measure of prediction: % of relative improvement using one to predict the other (= Kruskal-Wallis-λ) Two different methods in measuring the first guess (modal resp. observed probabilities)		
Eta	[0;1]	Nominal x Interval (ANOVA). Eta² is the percentage of variation of Y (interval) explained by X (nominal) by assuming linear relation between X and Y. Two eta values are computed: Nominal x Interval is interpretable.		
McNemar test		Typical test for repeated measures. Before after situation. (2 x 2)		
Risk		Odds Ratio and Relative Risk estimators		
Cochran's and Mantel-Haenszel statistics		Compare Odds Ratios of layers:		
		Cochran-Mantel-Haenszel Test:	$H_0 : \theta_1 = \theta_2 \ldots = \theta_k = 1$	conditional independent
		Breslow-Day Test	$H_0 : \theta_1 = \theta_2 \ldots = \theta_k$	homogeneous association
		Mantel-Haenszel Common Odds Ratio	Estimates one θ	One global Odds Ratio

Not Recommended: Overview

Fisher's Exact & Chi-Square & Likelihood ratio

Fisher's Exact test	Tests for differences in proportions. Exact test (SPSS: 2 x 2 ; SAS: r x c)
Pearson Chi-Square χ^2	Tests for differences in proportions. Asymptotic test. H_0: Independency. ($\leq$ 20% cells expected freq. < 5)
Likelihood ratio	f(observed/expected) via χ^2

Is myocardial infarct rate dependent from smoker status ?

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	9	3	12
	No	3	7	10

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	75%	25%	100%
	No	30%	70%	100%

P_1 = Proportion of Smokers having Myocardial infarct

P_2 = Proportion of Non-Smokers having Myocardial infarct

H_0: Myocardial infarct rate is **independent** from smoking $\Leftrightarrow$ $P_1 = P_2$

Pearson Chi-Square (*)	0.04
Likelihood Ratio	0.03
* Warning: cells (25.0%) have expected count less than 5.	
Fisher' Exact Test	0.08

Our data did not show a dependence between smoker status and myocardial infarct (p = 0.08 Fisher's-Exact)

a difference in rates of myocardial infarct between smokers and non-smokers

Guido Lüchters 01/04

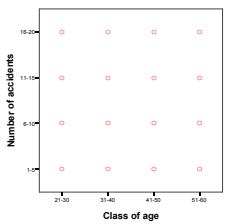
Spearman

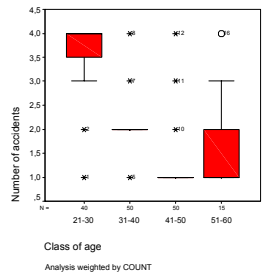
Spearman correlation ρ_s [-1;1] Monotone correlation. Invariant against any monotone transformation. Two ordinal variables.

Relation between age of driver and number of accidents

Class of age * Number of accidents Crosstabulation

			Number of accidents				Total
			1-5	6-10	11-15	16-20	
Class of age	21-30	Count	3	4	3	30	40
		% within Class of age	7,5%	10,0%	7,5%	75,0%	100,0%
	31-40	Count	4	40	3	3	50
		% within Class of age	8,0%	80,0%	6,0%	6,0%	100,0%
	41-50	Count	40	3	3	4	50
		% within Class of age	80,0%	6,0%	6,0%	8,0%	100,0%
	51-60	Count	10	2	2	1	15
		% within Class of age	66,7%	13,3%	13,3%	6,7%	100,0%
Total		Count	57	49	11	38	155
		% within Class of age	36,8%	31,6%	7,1%	24,5%	100,0%





Symmetric Measure	Value	Approx. Sig.
Interval by Interval Pearson's R	-,613	,000
Ordinal by Ordinal Spearman Correlation	-,730	,000

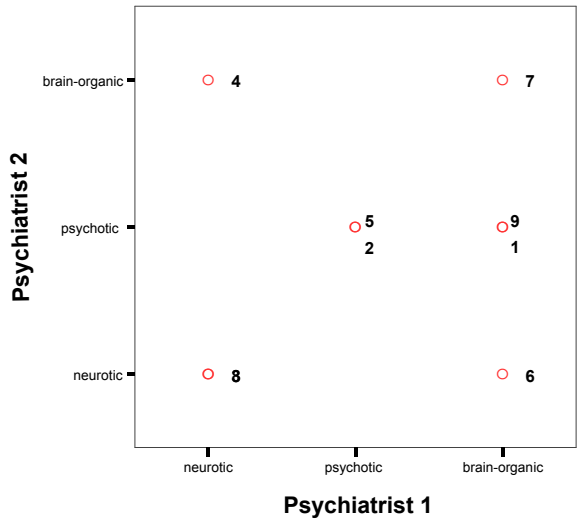
The older the driver the less accidents (p<0.000)
Assuming linear relation would explain 38% of accident variation
Assuming monotone relation would explain 53 % of accident variation.

Kappa Nominal

κ (Cohen's Kappa) [0;1] Measure of concordance or reliability between n observers (treatments) for nominal or ordinal data (n x n).

Diagnosis of 9 patients by 2 psychiatrics.

Patient	Psychiatrist 1	Psychiatrist 2
1	brain-organic	psychotic
2	psychotic	psychotic
3	neurotic	neurotic
4	neurotic	brain-organic
5	psychotic	psychotic
6	brain-organic	neurotic
7	brain-organic	brain-organic
8	neurotic	neurotic
9	brain-organic	psychotic



Psychiatrist 1 * Psychiatrist 2 Crosstabulation

Count

		Psychiatrist 2			
		neurotic	psychotic	brain-organic	Total
Psychiatrist 1	neurotic	2		1	3
	psychotic		2		2
	brain-organic	1	2	1	4
Total		3	4	2	9

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	,357	,224	1,664	,096
N of Valid Cases		9			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Our data did not show similar diagnosis' of 9 patients by two psychiatrics (p = 0.1)

The similarity (Kappa = 0.4) of diagnosis' of 9 patients by two psychiatrics could be due to chance (p = 0.1)

Kappa Ordinal

κ (Cohen's Kappa) [0;1] Measure of concordance or reliability between n observers (treatments) for nominal or ordinal data (n x n).

A company is interested in expanding its business, but has to choose from among 20 different sites to build. They have hired two consultants to separately evaluate these sites.

Site	Rating from consultant 1	Rating from consultant 2
1	Good	Good
2	Poor	Poor
3	Fair	Poor
4	Fair	Poor
5	Good	Good
6	Poor	Poor
7	Poor	Poor
8	Fair	Poor
9	Good	Good
10	Fair	Poor
11	Poor	Poor
12	Fair	Poor
13	Fair	Fair
14	Poor	Poor
15	Fair	Fair
16	Poor	Poor
17	Fair	Good
18	Good	Good
19	Good	Poor
20	Fair	Good

Kappa Ordinal

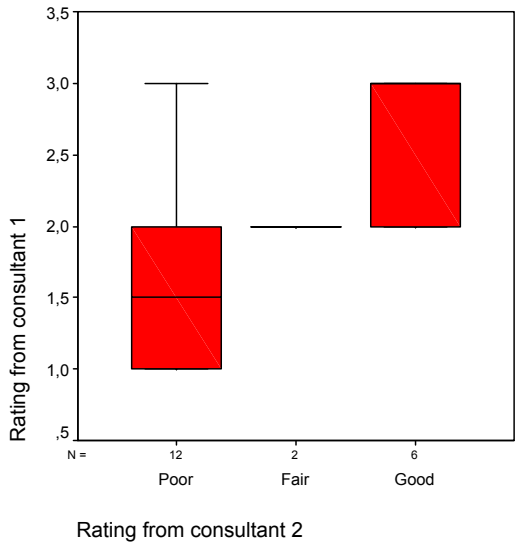
κ (Cohen's Kappa) [0;1] Measure of concordance or reliability between n observers (treatments) for nominal or ordinal data (n x n).

A company is interested in expanding its business, but has to choose from among 20 different sites to build. They have hired two consultants to separately evaluate these sites.

Rating from consultant 1 * Rating from consultant 2 Crosstabulation

Count

		Rating from consultant 2			
		Poor	Fair	Good	Total
Rating from consultant 1	Poor	6			6
	Fair	5	2	2	9
	Good	1		4	5
Total		12	2	6	20



Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	,610	,128	3,956	,000
Measure of Agreement	Kappa	,429	,131	3,333	,001
N of Valid Cases		20			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Consultants evaluate site similar (p =0.000) ⇔ The value of kappa is significantly (p=0.000) different from 0. Kappa = 0.4 suggests that the consultants' ratings are largely similar, with some exceptions.

Lambda (data)

λ and Goodman-Kruskal tau [0;1] Measure of prediction: % of relative improvement using one to predict the other (= Kruskal-Wallis-λ)
Two different methods in measuring the first guess (modal resp. observed probabilities)

A retail company conducted surveys of 1040 customers at 2 store locations to determine customer satisfaction rates

Store	Contact to employee	Service satisfaction	Customers
1	No	Somewhat Negative	50
1	No	Neutral	50
1	No	Somewhat Positive	45
1	No	Strongly Positive	40
1	Yes	Strongly Negative	40
1	Yes	Somewhat Negative	50
1	Yes	Neutral	50
1	Yes	Somewhat Positive	45
1	Yes	Strongly Positive	75
2	No	Strongly Negative	75
2	No	Somewhat Negative	40
2	No	Neutral	40
2	No	Somewhat Positive	45
2	No	Strongly Positive	50
2	Yes	Strongly Negative	50
2	Yes	Somewhat Negative	75
2	Yes	Neutral	75
2	Yes	Somewhat Positive	45
2	Yes	Strongly Positive	50

Lambda (global)

λ and Goodman-Kruskal tau [0;1] Measure of prediction: % of relative improvement using one to predict the other (= Kruskal-Wallis-λ)
Two different methods in measuring the first guess (modal resp. observed probabilities)

A retail company conducted surveys of 1040 customers at 2 store locations to determine customer satisfaction rates

Store * Service satisfaction Crosstabulation

			Service satisfaction					
			Strongly Negative	Somewhat Negative	Neutral	Somewhat Positive	Strongly Positive	Total
Store	Store 1	Count	100	100	90	115	115	520
		% within Store	19,2%	19,2%	17,3%	22,1%	22,1%	100,0%
	Store 2	Count	115	115	90	100	100	520
		% within Store	22,1%	22,1%	17,3%	19,2%	19,2%	100,0%
Total		Count	215	215	180	215	215	1040
		% within Store	20,7%	20,7%	17,3%	20,7%	20,7%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4,186[a]	4	,381
N of Valid Cases	1040		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 90,00.

Directional Measures

			Value	Approx. Sig.
Nominal by Nominal	Lambda	Symmetric	,033	,121
		Store Dependent	,058	,148
		Service satisfaction Dependent	,018	,306
	Goodman and Kruskal tau	Store Dependent	,004	,382[c]
		Service satisfaction Dependent	,001	,363[c]

c. Based on chi-square approximation

Lambda = 0.018 p = 0.306 :
To predict the service satisfaction of customers, the knowledge of store location improves pure guessing by around 2 % which is not a significant improvement.

Lambda (conditional)

λ and Goodman-Kruskal tau [0;1] Measure of prediction: % of relative improvement using one to predict the other (= Kruskal-Wallis-λ)
Two different methods in measuring the first guess (modal resp. observed probabilities)

A retail company conducted surveys of 1040 customers at 2 store locations to determine customer satisfaction rates

Store * Service satisfaction * Contact with employee Crosstabulation

Contact with employee				Service satisfaction					Total
				Strongly Negative	Somewhat Negative	Neutral	Somewhat Positive	Strongly Positive	
No	Store	Store 1	Count	50	50	45	40	40	225
			% within Store	22,2%	22,2%	20,0%	17,8%	17,8%	100,0%
		Store 2	Count	40	40	45	50	50	225
			% within Store	17,8%	17,8%	20,0%	22,2%	22,2%	100,0%
	Total		Count	90	90	90	90	90	450
			% within Store	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%
Yes	Store	Store 1	Count	50	50	45	75	75	295
			% within Store	16,9%	16,9%	15,3%	25,4%	25,4%	100,0%
		Store 2	Count	75	75	45	50	50	295
			% within Store	25,4%	25,4%	15,3%	16,9%	16,9%	100,0%
	Total		Count	125	125	90	125	125	590
			% within Store	21,2%	21,2%	15,3%	21,2%	21,2%	100,0%

Lambda (conditional)

λ and Goodman-Kruskal tau [0;1] Measure of prediction: % of relative improvement using one to predict the other (= Kruskal-Wallis-λ)
Two different methods in measuring the first guess (modal resp. observed probabilities)

A retail company conducted surveys of 1040 customers at 2 store locations to determine customer satisfaction rates

Chi-Square Tests

Contact with employee		Value	df	Asymp. Sig. (2-sided)
No	Pearson Chi-Square	4,444[a]	4	,349
	N of Valid Cases	450		
Yes	Pearson Chi-Square	20,000[a]	4	,000
	N of Valid Cases	590		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 45,00.

Directional Measures

Contact with employee				Value	Approx. Sig.
No	Nominal by Nominal	Lambda	Symmetric	,051	,118
			Store Dependent	,089	,135
			Service satisfaction Dependent	,028	,291
		Goodman and Kruskal tau	Store Dependent	,010	,350[c]
			Service satisfaction Dependent	,002	,350[c]
Yes	Nominal by Nominal	Lambda	Symmetric	,099	,001
			Store Dependent	,169	,001
			Service satisfaction Dependent	,054	,025
		Goodman and Kruskal tau	Store Dependent	,034	,001[c]
			Service satisfaction Dependent	,009	,000[c]

c. Based on chi-square approximation

Lambda = 0.054 p = 0.025 :
To predict the service satisfaction of customers who had contact with employees, the knowledge of store location improves pure guessing by around 5 % (p=.03)

Eta

Eta [0;1] Nominal x Interval (ANOVA). Eta² is the percentage of variation of Y (interval) explained by X (nominal) by assuming linear relation between X and Y.
Two eta values are computed: Nominal x Interval is interpretable.

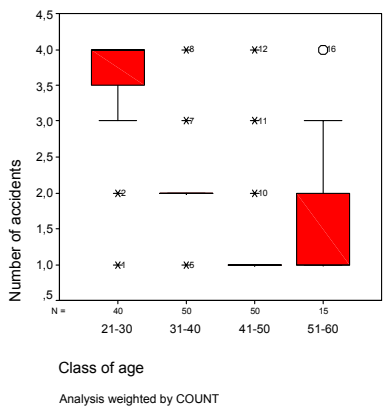
Relation between age of driver and number of accidents

Class of age * Number of accidents Crosstabulation

			Number of accidents				Total
			1-5	6-10	11-15	16-20	
Class of age	21-30	Count	3	4	3	30	40
		% within Class of age	7,5%	10,0%	7,5%	75,0%	100,0%
	31-40	Count	4	40	3	3	50
		% within Class of age	8,0%	80,0%	6,0%	6,0%	100,0%
	41-50	Count	40	3	3	4	50
		% within Class of age	80,0%	6,0%	6,0%	8,0%	100,0%
	51-60	Count	10	2	2	1	15
		% within Class of age	66,7%	13,3%	13,3%	6,7%	100,0%
Total		Count	57	49	11	38	155
		% within Class of age	36,8%	31,6%	7,1%	24,5%	100,0%

Directional Measures

			Value
Nominal by Interval	Eta	Class of age Dependent	,674
		Number of accidents Dependent	,697



49 % (*) of the variation of Number of accidents are explained by assuming that accidents are linear depending on age

(*) 0.697²

McNemar

McNemar test Typical test for repeated measures. Before after situation. (2 x 2)

Is gingival bleeding before and after a surgery depending on previous bleeding status or other reasons like the surgery

Gingival bleeding before * Gingival bleeding after Crosstabulation

			Gingival bleeding after		
			Yes	No	Total
Gingival bleeding before	Yes	Count	9	20	29
		% within Gingival bleeding before	31,0%	69,0%	100,0%
	No	Count	6	39	45
		% within Gingival bleeding before	13,3%	86,7%	100,0%
Total		Count	15	59	74
		% within Gingival bleeding before	20,3%	79,7%	100,0%

Chi-Square Tests

	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
McNemar Test	,009[a]	,005[a]	,003[a]

a. Binomial distribution used.

Gingival bleeding after surgery is not independent from gingival bleeding before surgery (p = 0.009)

Odds Ratio: Definition

		Myocardial infarct		
		Yes	No	
Smoker	Yes	50	50	100
	No	20	80	100

		Myocardial infarct		
		Yes	No	
Smoker	Yes	0.5	0.5	1
	No	0.2	0.8	1

		Myocardial infarct		
		Yes	No	
Smoker	Yes	P_S	$1-P_S$	1
	No	P_{NS}	$1-P_{NS}$	1

Odd for smoker to have an infarct = $\frac{\text{chance to have one}}{\text{chance not to have one}} = \frac{0.5}{0.5} = \mathbf{1}$

Odd for non smoker to have an infarct = $\frac{\text{chance to have one}}{\text{chance not to have one}} = \frac{0.2}{0.8} = \mathbf{0.25}$

$$\text{Odds Ratio} = \frac{\text{Odd for smoker}}{\text{Odd for non-smoker}} = \frac{1}{0.25} = 4$$

Odds Ratio = Cross-Product Ratio

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	0.5	0.5	1
	No	0.2	0.8	1

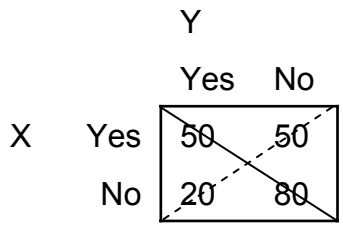
		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	50	50	100
	No	20	80	100

$$\text{Odds Ratio} = \frac{\text{Odd for smoker}}{\text{Odd for non-smoker}} = \frac{1}{0.25} = \frac{\frac{0.5}{0.5}}{\frac{0.2}{0.8}} = \frac{\frac{50}{50}}{\frac{20}{80}} = \frac{50 \bullet 80}{20 \bullet 50} = 4$$

$$\text{Odds Ratio} = \text{Cross-Product Ratio} = \frac{a \bullet d}{b \bullet c}$$

	Yes	No
Yes	a	b
No	c	d

Odds Ratio: Two Properties



θ (X,Y) := Odds Ratio (X,Y) < theta >

Two Properties:

θ (X,Y) = θ (Y,X) = θ

θ = 1 $\Leftrightarrow$ X and Y independent

Relative Risk

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	50	50	100
	No	20	80	100

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	0.5	0.5	1
	No	0.2	0.8	1

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	P_S	$1-P_S$	1
	No	P_{NS}	$1-P_{NS}$	1

Relative Risk of smokers to have an infarct = $\frac{\text{chance to have one for smokers}}{\text{chance have one for non smokers}} = \frac{0.5}{0.2}$ = **2.5**

ψ (X|Y=yes) := Relative Risk of X to face Y=yes < psi >

= ψ of column 1

= For cohort Myocardial infarct = yes SPSS

= Of cohort (Col 1 risk) SAS

Calculation of the Relative Risk is only allowed in prospective studies (Cohort)

Odds Ratio and Relative Risk

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	50	50	100
	No	20	80	100

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	0.5	0.5	1
	No	0.2	0.8	1

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	P_1	$1-P_1$	1
	No	P_2	$1-P_2$	1

Properties::

$$\theta = \frac{\psi \text{ (Column 1)}}{\psi \text{ (Column 2)}}$$

To check your calculation or output

$$\psi \text{ (Column 1)} = \theta \bullet \frac{(1-p_2)}{(1-p_1)}$$

If p_1 and p_2 is small then θ is a good proxy for ψ

Usually you can **not** estimate ψ in case-control studies (constructed margins).

You can **always** estimate θ - but this might lead to some false conclusions (see example)

Odds Ratio and Relative Risk: Output

		Myocardial infarct: Yes	Myocardial infarct: No	
Smoker	Yes	50	50	100
	No	20	80	100

SPSS 11

Risk Estimate

	Value	95% Confidence Interval: Lower	95% Confidence Interval: Upper
Odds Ratio for Smoker (Yes / No)	4,000	2,136	7,492
For cohort Myocardial Infarct = Yes	2,500	1,613	3,875
For cohort Myocardial Infarct = No	,625	,502	,778
N of Valid Cases	200		

Two Relative Risk estimates

SAS 8e

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	4.0000	2.1357	7.4916
Cohort (Col1 Risk)	2.5000	1.6129	3.8750
Cohort (Col2 Risk)	0.6250	0.5020	0.7781

Sample Size = 200

Two Relative Risk estimates

Pro:

One value for a table

Easy approach in Logistic Regression

Contra

Cheating is easy

- Small becomes big (medium and large samples)
- False conclusions (Cause and Effect Confusion)

Hard to interpret

Odds Ratio: Contra - Small becomes big

Mammography screening reduces the risk of dying of mamma carcinoma by 25% (*)

Medium sample size

		Dying within 10 years of mamma carcinoma: Yes	No	
Participate in regular screening	Yes	3	997	1,000
	No	4	996	1,000

Odd of dying for participating	= 3 / 997	= 0.003…	
Odd of dying for non participating	= 4 / 996	= 0.004…	CI
Odds Ratio	= 0.003… / 0.004…	= 0.749	[0.17 ; 3.36]
Relative Risk (Col 1)	= (4/1000) / (3/1000)	= 0.750	[0.17 ; 3.34]

(*) Data source: Mühlhauser, I., Informed Decision Making, Hamburg, 2001

Odds Ratio: Contra - Small becomes big

Mammography screening reduces the risk of dying of mamma carcinoma by 10-30% (*)

Large sample size

		Dying within 10 years of mamma carcinoma: Yes	Dying within 10 years of mamma carcinoma: No	
Participate in regular screening	Yes	290	99,710	100,000
	No	360	99,640	100,000

		CI
Odds Ratio	= 0.8	[0.7 ; 0.9]
Relative Risk (Col 1)	= 0.8	[0.7 ; 0.9]

The big number blurs (**):

- 30% does not mean 30 of 100 women rather it is a quotient of quotients and hard to interpret
- as the prevalence of breast cancer is small (0.8%) for 100.000
  7.000 of 100.000 have to suffer from a false positive result (false positive rate = 7%)
  100 of 100.000 will have a false negative result (sensitivity = 90%)
  the individual chance of having breast cancer after a positive diagnosis is around 9%

(*) Data source: Mühlhauser, I., Information zum Mammographiescreening – vom Trugschluss zur Ent-Täuschung, Radiologe, 42, 2002

(**) Gigerenzer, G., "Calculated Risks", New York, 2002

Odds Ratio: Contra – False conclusions

262 young and middle-aged women (age<69) admitted to units with acute Myocardial Infarct (MI)

524 matched cases **were chosen** who were admitted to the same unit with other acute disorders (*)

	Smoker Yes	Smoker No	
MI	172	90	262
Control	175	349	524

Case-Control study

row margin constructed

Odd for smokers under MI	= 172 / 90	= 1.91
Odd for non- smoker under Control	= 175 / 349	= 0.50
θ = Odds Ratio of smoking (MI / Control)	= 1.91 / 0.5 = (172●349)/(90●175)	= 3.81

We would like to have the Relative Risk for smokers to have an infarct

But that would demand a Cohort study.

Instead of doing so, we just revert the table.

		MI	Control	
Smoker	Yes	172	175	347
	No	90	349	439

Odd for MI under smokers	= 172 / 175	= 0.983
Odd for MI under non smoker	= 90 / 349 = (172●349)/(90●175)	= 0.258
θ = Odds Ratio of MI (smoker / non-smoker)	= 0.983 / 0.258	= 3.81

Because of $\theta(X,Y) = \theta(Y,X) = \theta$ we calculate θ from table 1 and apply it to table 2

Because of Relative Risk ~ θ we then conclude the **false** Relative Risk (of table 2)

The relative risk for smoker to suffer a MI is 4 times higher then for non-smokers

This is a wrong conclusion, as we did not perform a cohort trial but just used two algebraic properties of the odds ratio.

(*) Data source slightly modified, Agresti, A., "An Introduction to Categorial Data Analysis, 1997

Meta Analysis
Odds Ratios equal

Bonn

		Result: success	Result: failure	
Therapy	A	20	10	30
	B	10	20	30

Odd of therapy	A	20 / 10	**2**	
Odd of therapy	B	10 / 20	**0.5**	
Odds Ratio	A/B	2 / 0.5	**4**	[1.4 ; 11.7]

Frankfurt

		Result: success	Result: failure	
Therapy	A	2000	1000	3000
	B	1000	2000	3000

Odd of therapy	A	2000 / 1000	**2**	
Odd of therapy	B	1000 / 2000	**0.5**	
Odds Ratio	A/B	2 / 0.5	**4**	[3.6 ; 4.5]

Cochran – Mantel – Haenszel Test — $H_0: OR_1 = 1$ and $OR_2 = 1$ — There is no significant OR (conditionally independent)

p = 0.000 **high significant** Result is not independent in at least one clinic (here in both)

Breslow – Day Test — $H_0: OR_1 = OR_2$ — All ORs are equal (associations homogeneous)

p = 1 **not significant** We could not find a difference between the two Odds Ratios

All

		Result: success	Result: failure	
Therapy	A	2020	1010	3030
	B	1010	2020	3030

Odd of therapy	A	2020 / 1010	**2**	
Odd of therapy	B	1010 / 2020	**0.5**	
Odds Ratio	A/B	2 / 0.5	**4**	[3.6 ; 4.5]

Mantel-Haenszel Common Odds Ratio Estimate **4** [3.6 ; 4.5]

Guido Lüchters 01/04

Meta Analysis
Odds Ratios opposite

Bonn

		Result: success	Result: failure	
Therapy	A	200	100	300
	B	100	200	300

Odd of therapy	A	200 / 100	**2**	
Odd of therapy	B	100 / 200	**0.5**	
Odds Ratio	A/B	2 / 0.5	**4**	[2.8 ; 5.6]

Frankfurt

		Result: success	Result: failure	
Therapy	A	100	200	300
	B	200	100	300

Odd of therapy	A	100 / 200	**0.5**	
Odd of therapy	B	200 / 100	**2**	
Odds Ratio	A/B	0.5 / 2	**0.25**	[0.2 ; 0.4]

Cochran – Mantel – Haenszel Test $H_0: OR_1 = 1$ and $OR_2 = 1$ — **There is no significant OR (conditionally independent)**

p = 0.9 **not significant** known weakness of this test

Breslow – Day Test $H_0: OR_1 = OR_2$ — **All ORs are equal (associations homogeneous)**

p = 0.000 **high significant** The two odds ratios are not equal

All

		Result: success	Result: failure	
Therapy	A	300	300	600
	B	300	300	600

Odd of therapy	A	300 / 300	**1**	
Odd of therapy	B	300 / 300	**1**	
Odds Ratio	A/B	1 / 1	**1**	[0.8 ; 1.3]

Mantel-Haenszel Common Odds Ratio Estimate **1** [0.8 ; 1.2]

Guido Lüchters 01/04

Meta Analysis
Simpson Paradox Overview

Bonn

		Result: success	Result: failure	∑	% success
Therapy	A	900	100	1000	90%
	B	1000	1000	2000	50%

Odds Ratio A/B = **9**

Frankfurt

		Result: success	Result: failure	∑	% success
Therapy	A	300	2700	3000	10%
	B	30	970	1000	3%

Odds Ratio A/B = **3.6**

All

		Result: success	Result: failure	∑	% success
Therapy	A	1200	2800	4000	30%
	B	1030	1970	3000	34%

Odds Ratio A/B = **0.8**

Guido Lüchters 01/04

Meta Analysis
Simpson Paradox large sample sizes

Bonn

		Result: success	Result: failure		% success
Therapy	A	900	100	1000	90
	B	1000	1000	2000	50

Odd of therapy	A	900 / 100	**9**	
Odd of therapy	B	1000 / 1000	**1**	
Odds Ratio	A/B	9 / 1	**9**	[7.2 ; 11.3]

Frankfurt

		Result: success	Result: failure		% success
Therapy	A	300	2700	3000	10
	B	30	970	1000	3

Odd of therapy	A	300 / 2700	**0.11**	
Odd of therapy	B	30 / 970	**0.03**	
Odds Ratio	A/B	0.1 / 0.01	**3.6**	[2.5 ; 5.3]

Cochran – Mantel – Haenszel Test — $H_0: OR_1 = 1$ and $OR_2 = 1$ — **There is no significant OR (conditionally independent)**

p = 0.000 — **high significant** — Result is not independent in at least one clinic (here both)

Breslow – Day Test — $H_0: OR_1 = OR_2$ — **All ORs are equal (associations homogeneous)**

p = 0.000 — **high significant** — Odds Ratios are not equal

All

		Result: success	Result: failure		% success
Therapy	A	1200	2800	4000	30
	B	1030	1970	3000	34

Odd of therapy	A	1200 / 2800	**0.4**	
Odd of therapy	B	1030 / 1970	**0.5**	
Odds Ratio	A/B	2 / 0.5	**0.8**	[0.7 ; 0.9]

Mantel-Haenszel Common Odds Ratio Estimate — **7** — [5.7 ; 8.5]

Guido Lüchters 01/04

Meta Analysis
Simpson Paradox medium sample sizes

Bonn

		Result: success	Result: failure		% success
Therapy	A	90	10	100	90
	B	100	100	200	50

Odd of therapy	A	90 / 10	**9**	
Odd of therapy	B	100 / 100	**1**	
Odds Ratio	A/B	9 / 1	**9**	[4 ; 18]

Frankfurt

		Result: success	Result: failure		% success
Therapy	A	30	270	300	10
	B	3	97	100	3

Odd of therapy	A	30 / 270	**0.11**	
Odd of therapy	B	3 / 97	**0.03**	
Odds Ratio	A/B	0.1 / 0.01	**3.6**	[1.1 ; 12]

Cochran – Mantel – Haenszel Test — $H_0: OR_1 = 1$ and $OR_2 = 1$ — **There is no significant OR (conditionally independent)**

p = 0.000 — **high significant** — Result is not independent in at least one clinic (here both)

Breslow – Day Test — $H_0: OR_1 = OR_2$ — **All ORs are equal (associations homogeneous)**

p = 0.19 — **not significant** — We could not find a difference between the two Odds Ratios

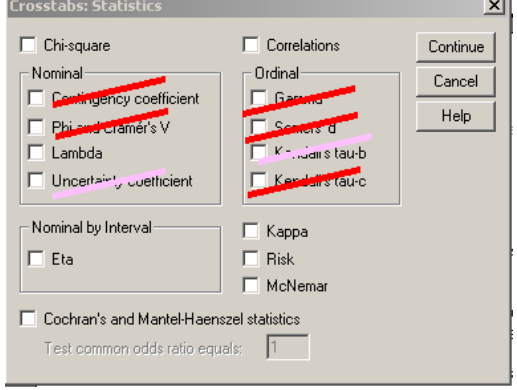
All

		Result: success	Result: failure		% success
Therapy	A	120	280	400	30
	B	103	197	300	34

Odd of therapy	A	120 / 280	**0.4**	
Odd of therapy	B	103 / 197	**0.5**	
Odds Ratio	A/B	2 / 0.5	**0.8**	[0.6 ; 1.1]

Mantel-Haenszel Common Odds Ratio Estimate — **7** — [4 ; 13]

Guido Lüchters 01/04



Not Recommended: Overview

Be careful		
Pearson correlation ρ	[-1;1]	Linear (!) correlation. Two ordinal variables.
τ_b	[-1;1]	Measure of association between two ordinal variables via concordant pairs corrected for ties. Two ordinal variables. No advantage over Spearman.
Linear-by-Linear Association = Cochran's Mantel-Haenszel **Chi-Square**		**SPSS** SAS = $(n-1)\rho^2$ Tests for linear (!) association. Two ordinal variables.
uncertainty coefficient	[0;1]	rarely used, based on calculation of entropy
Not recommended		
Contingency coefficient	[0;1)	= $f(\chi^2) < 1$ depending on number of rows and columns
φ	[-1;1) [0;1)	= $f(\chi^2)$ [-1;1) for 2 x 2 and [0;1) for r x c depending on marginal distributions.
γ	[-1;1]	$\approx \tau_b$ but not corrected for ties
Somer's d	[-1;1]	$\approx \tau_b$ but asymmetric i.e. row is regarded as independent. Corrects ties only for rows.
τ_c	(-1;1)	$\approx \tau_b$ but not corrected for ties -1 or +1 only for n x n tables

End Categorial

Thank you for your attention