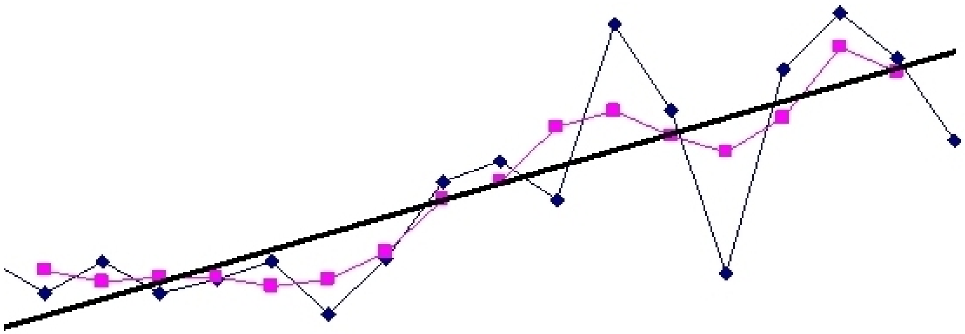
Author: Tibor Nyári PhD

University of Szeged
Department of Medical Physics and Informatics

www.model.u-szeged.hu
www.szote.u-szeged.hu/dmi

Kappa, ODDS ratio, Relative risks: Measurement of agreement



Cohen's Kappa

- Kappa measures the **agreement** between two test results.
  - Jacob Cohen (1923 – 1998) was a US statistician and psychologist.
  - He described kappa statistic in 1960.
- H_0: $\kappa=0$
- H_A: $\kappa\neq0$

Measuring agreements (observed frequencies)

	Test 1			
Test 2	Positive	Negative	Total	
Positive	**a**	b	Z_1=a+b	Z_1/N
Negative	c	**d**	Z_2=c+d	Z_2/N
Total	S_1=a+c	S_2=b+d	N	N
	S_1/N	S_2/N		

- Agreement in the diagonal.
- Probability of a positive and negative results of the Test I are S_1/N and S_2/N, respectively
- Probability of a positive and negative results of the Test II are : Z_1/N and Z_2/N, respectively
- Observed probability of agreement: p_{obs}=(a+d)/N

$$p_O = \frac{a + d}{N}$$

Expected frequencies

$$P(AB) = P(A)P(B) \Rightarrow \frac{E_{11}}{N} = \frac{S_1}{N}\frac{Z_1}{N}$$

	Test I	
	Positiv	Negativ
Positiv	**E11** $= \frac{S_1}{N}\frac{Z_1}{N}N$	E12
Negativ	E21	**E22** $= \frac{S_2}{N}\frac{Z_2}{N}N$

- Expected probability of agreement : $p_{Expected} = (E_{11} + E_{22})/N$

$$p_E = \frac{E_{11} + E_{22}}{N}$$

Cohen's kappa

$$p_{observed} = \frac{a + d}{N} \qquad p_E = \frac{E_{11} + E_{22}}{N}$$

$$\kappa = \frac{p_{Observed} - p_{Expected}}{1 - p_{Expected}}$$

Standard error (SE) for kappa:

$$\hat{se}(\kappa) = \sqrt{\frac{1}{(1 - p_E)^2 N}\left(p_E^2 + p_E - \sum_{i=1}^{l} \frac{S_i Z_i}{N}\{S_i + Z_i\}\right)}$$

The test statistic for kappa: $\left(\frac{\kappa}{\hat{se}(\kappa)}\right)^2$

This follows a χ^2 with 1 df.

$\chi^2_{table(\alpha=0,05;\ FG=1)}$–value = 3.841 (=1.96²)

Characteristics of kappa

- It takes the value 1 if the agreement is perfect and 0 if the amount of agreement is entirely attributable to chance.
- If $\kappa<0$ then the amount of agreement is less then would be expected by chance.
- If $\kappa>1$ then there is more than chance agreement.
- According to Fleiss:
  - Excellent agreement if $\kappa>0.75$
  - Good agreement if $0.4<\kappa<0.75$
  - Poor agreement if $\kappa<0.4$

Altman DG, Bland JM. Statistics Notes: Diagnostic tests : sensitivity and specificity *BMJ* 1994; 308 : 1552

- **Relation between results of liver scan and correct diagnosis**

Liver scan	Pathology		
	abnormal (+)	normal (-)	Total
abnormal (+)	231	32	263
normal(-)	27	54	81
Total	258	86	344

The expected freqencies

$$P(AB) = P(A)P(B) \Rightarrow \frac{E_{11}}{N} = \frac{S_1}{N}\frac{Z_1}{N}$$

- E_{11}=(263/344)*(258/344)*344=197.25
- E_{22}=(81/344)*(86/344)*344=20.25

Liver scan	Pathology		
	abnormal (+)	normal (-)	Total
abnormal (+)	**197.25**		263
normal(-)		**20.25**	81
Total	258	86	344

Cohen's kappa

- The observed p_{Obs} and p_{Exp} values are 0.828 and 0.63, respectively . Cohen's kappa (κ)=0.53.

$$p_{obs} = \frac{a + d}{N} = \frac{231 + 54}{344} = 0.828$$

$$p_E = \frac{E_{11} + E_{22}}{N} = \frac{197.25 + 20.25}{344} = 0.63$$

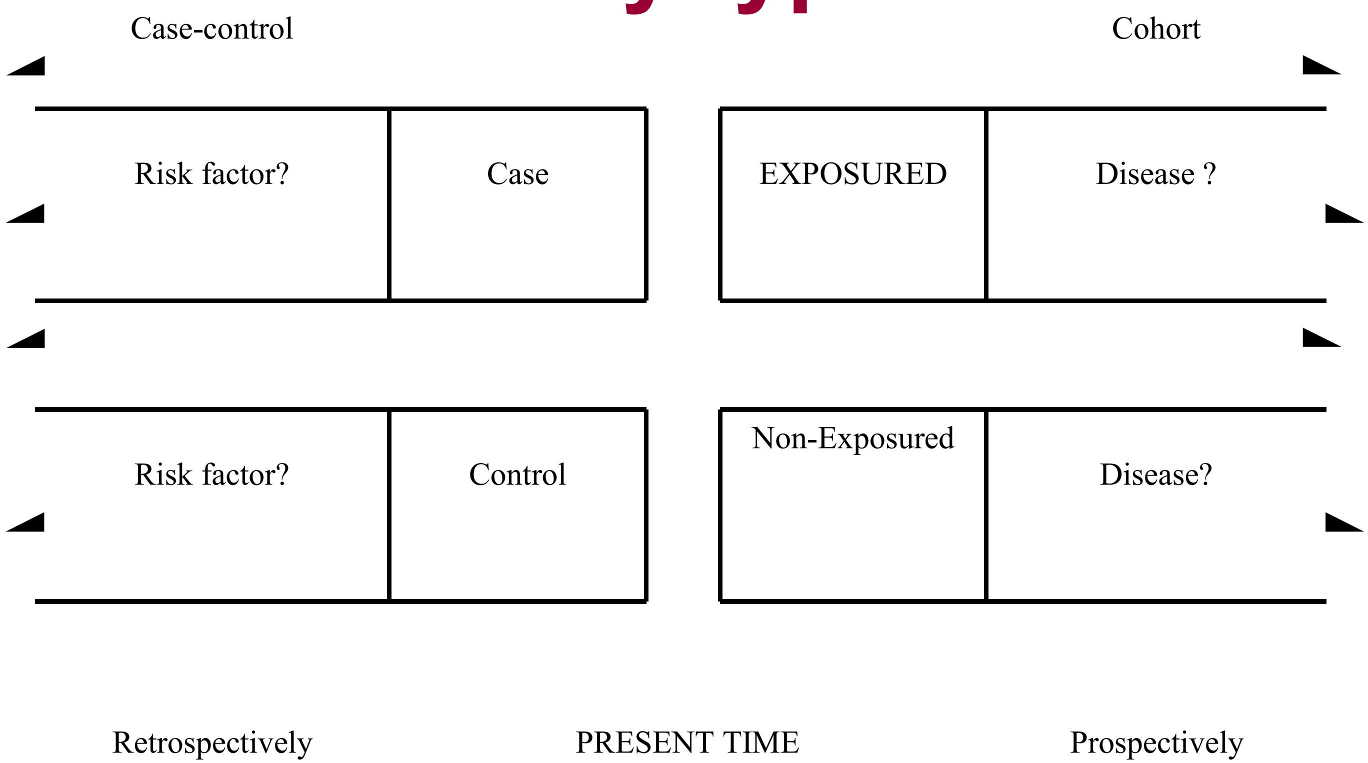
$$\kappa = \frac{p_{obs} - p_E}{1 - p_E} = \frac{0.828 - 0.632}{1 - 0.632} = 0.53$$

Decision

- Here $\kappa=0.53$
- As $0.4<\kappa\leq 0.75$: good agreement

Other applications

Study types


Case-control
Cohort
Risk factor?
Case
EXPOSURED
Disease ?
Risk factor?
Control
Non-Exposured
Disease?
Retrospectively
PRESENT TIME
Prospectively

Prevalence and incidence

- **Prevalence** quantifies the proportion of individuals in a population who have a specific disease at a specific point of time.

$$\text{Prevalence} = \frac{\text{number of existing cases of disease}}{\text{total population}} \quad \text{at a given time point}$$

- In contrast with the prevalence, the incidence quantifies the number of new events or cases of disease that develop in a population of individuals at risk during a specified period of time.

$$\text{Incidence risk} = \frac{\text{number of new cases of disease during a given period of time}}{\text{number at risk of contracting the disease at the beginning of the period}}$$

- There are two specific types of incidence measures: **incidence risk** and **incidence rate**.
  - The incidence risk is the proportion of people who become diseased during a specified period of time, and is calculated as

Odds ratio

- It measures of association in case-control studies.
- H_0: OR=1
- H_A: OR≠1

$$OR = \frac{a/b}{c/d} = \frac{ad}{cb} \text{ and } SE(OR) = \sqrt{\left(\frac{1}{a}\right) + \left(\frac{1}{b}\right) + \left(\frac{1}{c}\right) + \left(\frac{1}{d}\right)}$$

- An alternative measure of incidence is the odds of disease to non-disease. This equals the total number of cases divided by those still at risk at the end of the study. Using the notation of previous Table , reproduced on next slide:

Odds ratio

	Disease		
	Yes	No	Total
Exposed	a	b	e=a+b
Non-exposed	c	d	f=c+d
Total	g=a+c	h=b+d	n=g+h

the odds of disease among the exposed is a/b and that among the unexposed is c/d.
Their ratio, called the odds ratio, is

$$OR = \frac{a/b}{c/d} = \frac{ad}{cb} \text{ and } SE(OR) = \sqrt{\left(\frac{1}{a}\right) + \left(\frac{1}{b}\right) + \left(\frac{1}{c}\right) + \left(\frac{1}{d}\right)}$$

Case-control studies

- In a case-control study, the sampling is carried out according to the disease rather than the exposure status.
- A group of individuals identified as having the disease, the cases, is compared with a group of individuals not having the disease, the controls, with respect to their prior exposure to the factor of interest.
- No information is obtained directly about the incidence in the exposed and non-exposed populations, and so the relative risk cannot be estimated; instead, the odds ratio is used as the measure of association.
- It can be shown, however, that for a rare disease the odds ratio is numerically equivalent to the relative risk.
- The 95% confidence interval for the odds ratio is calculated in the same way as that for relative risk:

$$95\%\ \text{CI} = e^{\left(\ln(OR) \pm 1.96\sqrt{\left(\frac{1}{a}\right)+\left(\frac{1}{b}\right)+\left(\frac{1}{c}\right)+\left(\frac{1}{d}\right)}\right)}, \text{ where } e = 2.718$$

Example

- The risk of HPV infection for smokers was measured in a study.
- H_0: OR=1
- H_A: OR≠1
- Calculate the odds ratio and 95% confidence interval using the data table

		HPV		
		Yes	No	Total
Smoking	Yes	**33**	**81**	114
	No	**58**	**225**	283
Total		91	306	397

$$OR = \frac{ad}{cb} = \frac{33*225}{81*58} = 1.58046$$

$$SE(OR) = \sqrt{\left(\frac{1}{33}\right)+\left(\frac{1}{225}\right)+\left(\frac{1}{81}\right)+\left(\frac{1}{58}\right)} = 0.25364$$

Results of Risk Estimate

$$OR = \frac{ad}{cb} = \frac{33 * 225}{81 * 58} = 1.58046$$

$$SE(OR) = \sqrt{\left(\frac{1}{33}\right)+\left(\frac{1}{225}\right)+\left(\frac{1}{81}\right)+\left(\frac{1}{58}\right)} = 0.25364$$

$$95\%\,\mathrm{CI} = 2.718^{\left(\ln(1.5804) \pm 1.96\sqrt{\left(\frac{1}{33}\right)+\left(\frac{1}{225}\right)+\left(\frac{1}{81}\right)+\left(\frac{1}{58}\right)}\right)} = 0.961\ ; 2.598$$

As OR=1.58 and its 95% confidence interval (95%CI) [0.96 – 2.59] contains 1, the H_0 is accepted.

SPSS results fo Risk Estimate

- As OR=1.58 and its 95% confidence interval (95%CI) [0.96 – 2.59] contains 1, the H_0 is accepted.

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for row (1,00 / 2,00)	1,580	,961	2,598
For cohort column = 1,00	1,412	,978	2,041
For cohort column = 2,00	,894	,784	1,019
N of Valid Cases	397		

Example

Research Report

European Addiction Research

Eur Addict Res 2005;11:38–43
DOI: 10.1159/000081415

Addictive Behaviour of Adolescents in Secondary Schools in Hungary

Table 2. Results of the univariate analysis in the ever-smoked and regular-smoker groups

		Children	Drug users	OR (95% CI)	p value
Ever-smoked					
Drug usage in the family	Yes	296	33	5.7 (1.7–19.0)	0.005
	No	23	9	1.0	
Living in a block of flat	Yes	71	14	1.8 (0.9–3.7)	0.086
	No	263	31	1.0	
Age, years	17–18	107	23	2.3 (1.2–4.6)	0.014
	15–16	171	18		
Sociable delinquencies	Yes	129	28	3.4 (1.7–6.7)	<0.001
	No	186	14	1.0	
School performance	Poor	17	6	15.0 (2.7–84.5)	0.002
	Acceptable	117	17	4.8 (1.0–21.0)	0.044
	Good	144	20	4.4 (1.0–19.7)	0.050
	Very good	57	2	1.0	
Truancy from school	Yes	50	13	3.3 (1.5–7.3)	0.003
	No	210	20	1.0	

SPSS Results

row * column Crosstabulation

Count

		column		Total
		1,00	2,00	
row	1,00	13	37	50
	2,00	20	190	210
Total		33	227	260

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for row (1,00 / 2,00)	3,338	1,527	7,296
For cohort column = 1,00	2,730	1,459	5,108
For cohort column = 2,00	,818	,690	,970
N of Valid Cases	260		

Results

- H_0: OR=1
- H_A: OR≠1

row * column Crosstabulation

Count

		column		Total
		1,00	2,00	
row	1,00	13	37	50
	2,00	20	190	210
Total		33	227	260

$$SE(OR) = \sqrt{\left(\frac{1}{13}\right)+\left(\frac{1}{37}\right)+\left(\frac{1}{20}\right)+\left(\frac{1}{190}\right)} = 0.399$$

- OR=(13*190)/ (37*20)=3.337 $\Rightarrow$ ln(OR)=1.205
- SE=0.399
- Lower bound =exp(1.205–1.96*0.399)=1.5269
- Upper bound =exp(1.205+1.96*0.399)=7.296
- As the 95% confidence interval (95%CI) [1.53 – 7.29] does not contain 1, thus H_A is accepted .

Mantel – Haenszel Odds ratio

	Risk yes	Risk no	Total	
1st group	$\boldsymbol{n_{111}}$	$\boldsymbol{n_{112}}$	n_{11+}	$p_{11}= n_{111} /n_{11+}$
2nd group	$\boldsymbol{n_{121}}$	$\boldsymbol{n_{122}}$	n_{12+}	$p_{12}= n_{121} /n_{12+}$
Total	n_{1+1}	n_{1+2}	n_1	

	Risk yes	Risk no	Total	
1st group	$\boldsymbol{n_{211}}$	$\boldsymbol{n_{212}}$	n_{21+}	$p_{21}= n_{211} /n_{21+}$
2nd group	$\boldsymbol{n_{221}}$	$\boldsymbol{n_{222}}$	n_{22+}	$p_{22}= n_{221} /n_{22+}$
Total	n_{2+1}	n_{2+2}	n_2	

$$EH = \frac{\sum_{i=1}^{2} \frac{n_{i11} * n_{i22}}{n_i}}{\sum_{i=1}^{2} \frac{n_{i12} * n_{i21}}{n_i}}$$

Example

- In a study the risk of coronary heart disease was investigated using ECG diagnosis by gender.

ecg * CHD * gender Crosstabulation

Count

gender			CHD: CHD_No	CHD: CHD_Yes	Total
Female	ecg	normal	11	4	15
		abnormal	10	8	18
	Total		21	12	33
Male	ecg	normal	9	9	18
		abnormal	6	21	27
	Total		15	30	45

- Female OR=2.2

Risk Estimate

	Value	95% Confidence Interval: Lower	95% Confidence Interval: Upper
Odds Ratio for row (1,00 / 2,00)	2,200	,504	9,611
For cohort column = 1,00	1,320	,790	2,206
For cohort column = 2,00	,600	,224	1,607
N of Valid Cases	33		

- Male OR=3.5

Risk Estimate

	Value	95% Confidence Interval: Lower	95% Confidence Interval: Upper
Odds Ratio for row (1,00 / 2,00)	3,500	,959	12,778
For cohort column = 1,00	2,250	,968	5,230
For cohort column = 2,00	,643	,388	1,064
N of Valid Cases	45		

Results

ecg * CHD * gender Crosstabulation

Count

gender			CHD: CHD_No	CHD: CHD_Yes	Total
Female	ecg	normal	11	4	15
		abnormal	10	8	18
	Total		21	12	33
Male	ecg	normal	9	9	18
		abnormal	6	21	27
	Total		15	30	45

$$EH = \frac{\sum_{i=1}^{2} \frac{n_{i11} * n_{i22}}{n_i}}{\sum_{i=1}^{2} \frac{n_{i12} * n_{i21}}{n_i}} =$$

$$EH = \frac{\frac{11 \cdot 8}{33} + \frac{9 \cdot 21}{45}}{\frac{10 \cdot 4}{33} + \frac{9 \cdot 6}{45}} = \frac{\frac{88}{33} + \frac{189}{45}}{\frac{40}{33} + \frac{54}{45}} = 2.84673$$

Mantel-Haenszel Common Odds Ratio Estimate

Estimate			2,847
ln(Estimate)			1,046
Std. Error of ln(Estimate)			,496
Asymp. Sig. (2-sided)			,035
Asymp. 95% Confidence Interval	Common Odds Ratio	Lower Bound	1,077
		Upper Bound	7,528
	ln(Common Odds Ratio)	Lower Bound	,074
		Upper Bound	2,019

The Mantel-Haenszel common odds ratio estimate is asymptotically normally distributed under the common odds ratio of 1,000 assumption. So is the natural log of the estimate.

Incidence risk

- The incidence risk, then, provides an estimate of the probability, or risk, that an individual will develop a disease during a specified period of time. This assumes that the entire population has been followed for the specified time interval for the development of the outcome under investigation. However, there are often varying times of entering or leaving a study and the length of the follow-up is not the same for each individual. The incidence rate utilizes information on the follow-up time for each subjects, and is calculated as

$$\text{Incidence rate} = \frac{\text{number of new cases of disease during a given period of time}}{\text{total "person - time" of observation}}$$

  - (The denominator is the sum of each individual's time at risk)

Example

- In a study of oral contraceptive (OC) use and bacteriuria, a total of 2 390 women aged between 16 to 49 years were identified who were free from bacteriuria. Of these, 482 were OC users at the initial survey in 1993. At a second survey in 1996, 27 of the OC users had developed bacteriuria. Thus,
- Incidence risk=27 per 482, or 5.6 percent during this 3-year period

Example

- In a study on postmenopausal hormone use and the risk of coronary heart disease, 90 cases were diagnosed among 32 317 postmenopausal women during a total of 105 782.2 person-years of follow-up. Thus,
- Incidence rate=90 per 105 782.2 person-years, or 85.1 per 1 000 000 person-years

Issues in the calculation of measures of incidence

- Precise definition of the denominator is essential.
- The denominator should, in theory, include only those who are considered at risk of developing the disease, i.e. the total population from which new cases could arise.
- Consequently, those who currently have or have already had the disease under study, or those who cannot develop the disease for reasons such as age, immunizations or prior removal of an organ, should, in principal, be excluded from the denominator.

Measures of association in cohort studies

	Lung cancer			
	Yes	No	Total	Incidence rate
Smokers	39	29 961	30 000	1.30/1000/year
Non-smokers	6	59 994	60 000	0.10/1000/year
Total	45	89 555	90 000	

Relative risk

	Disease		
	Yes	No	Total
Exposed	a	b	e=a+b
Non-exposed	c	d	f=c+d
Total	g=a+c	h=b+d	n=g+h

$$RR = \frac{I_{\exp}}{I_{non\ \exp}} = \frac{a/e}{c/f}$$

Relative risk

- The further the relative risk is from 1, the stronger the association.
- Its statistical association can be tested by using a 2 x 2 $\chi 2$ – test
- Confidence interval for RR:

$$95\%\ \text{CI} = \text{RR}^{\left(1 \pm 1.96\sqrt{\chi^{2}}\right)}$$

- In the above example, $95\%\ \text{CI} = 13.0^{\left(1 \pm 1.96\sqrt{55.5}\right)} = 6.7, 25.3$. The 95% confidence interval for the relative risk is therefore 6.7 to 25.3

Incidence rates (IR)

- Neuroblastoma is one of the most common solid tumour in children and the most common tumour in infants, accounting for about 9% of all cases of paediatric cancer and is a major contributor to childhood cancer mortality worldwide
- The incidence and distribution of the age and stage of neuroblastoma at diagnosis, and outcome in Hungary over a period of 11 years were investigated and compared with that reported for some Western European countries.

Age-specific and directly age-standardized (world population) incidence rates (per million) for neuroblastoma in Hungary (1988-1998) and in Austria (1987-1991)

		Hungary			Austria
Age-specific	IR	95%CI		IR	95%CI
< 1 year	60.9	(40.6-81.1)		65.8	(44.1-94.5)
1-4 years	25.5	(19.8-31.2)		17.0	(11.4-24.2)
5-9 years	4.2	(2.6-5.8)		3.1	(1.2-6.4)
10-14 years	1.7	(0.8-2.4)		1.3	(0.3-3.9)
Age-standardized	14.4	(12.6-16.2)		11.7	(9.0-14.5)