Methodology and Statistics for Linguistic Research
(LTR002M10, 2010/2011)

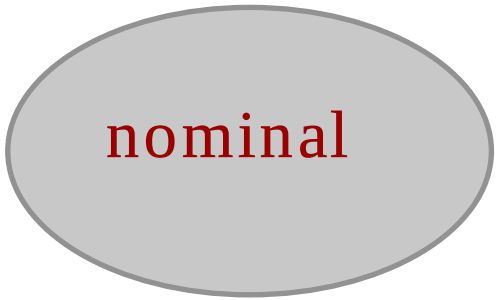
Cochran's Q test

Jet Vonk

Content

- Types of data
- Cochran's Q test
  - Who?
  - What?
  - When?
  - How?
  - Application?
- Cochran's Q test: Example
- Limitations of the Cochran's Q test
- Discussion
- References

Types of data



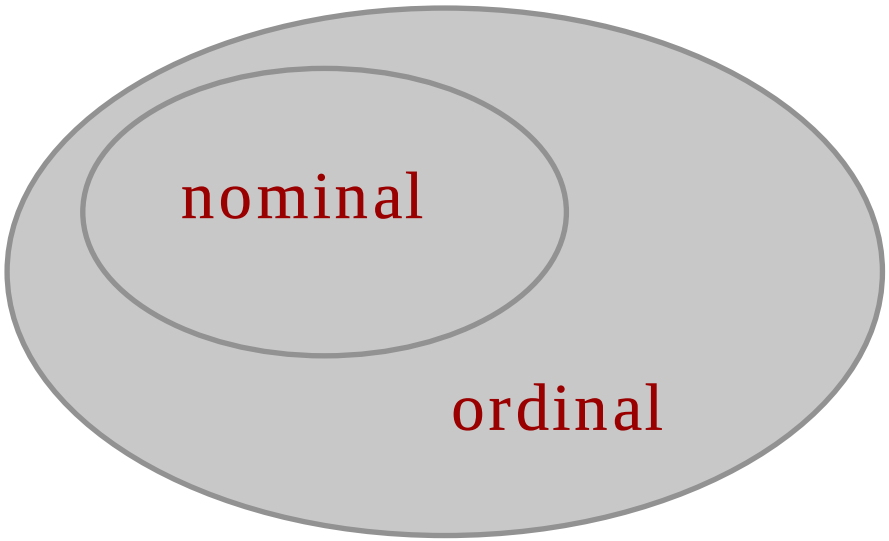
Types of data

› Nominal data: no specific order, no numerical meaning, differentiated by a naming system

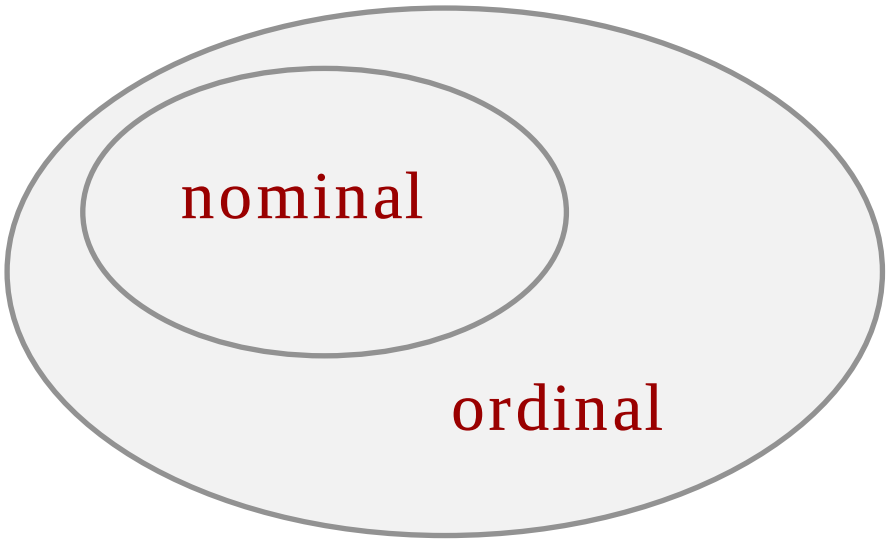
Examples:

- Men/women
- Set of countries

Types of data



Types of data

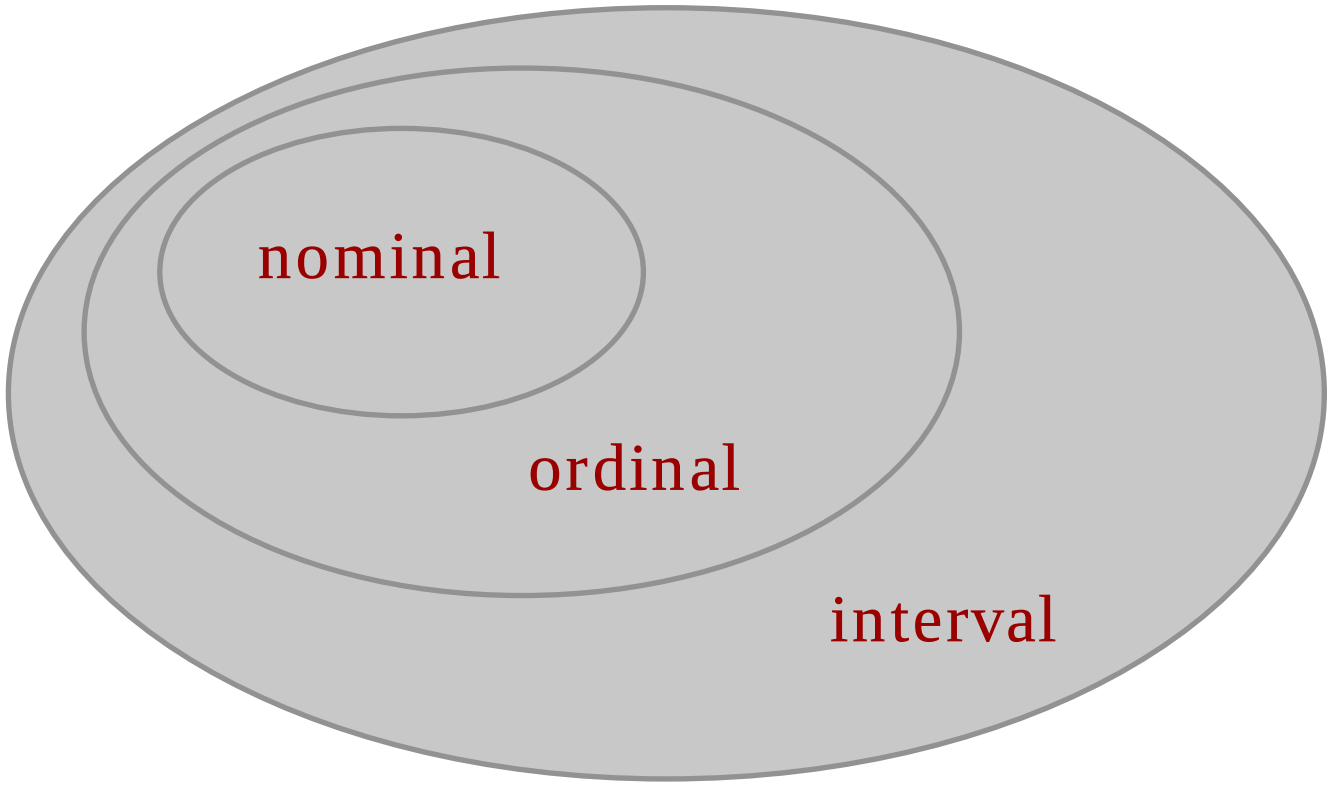


› Ordinal data: items have a specific order on the scale

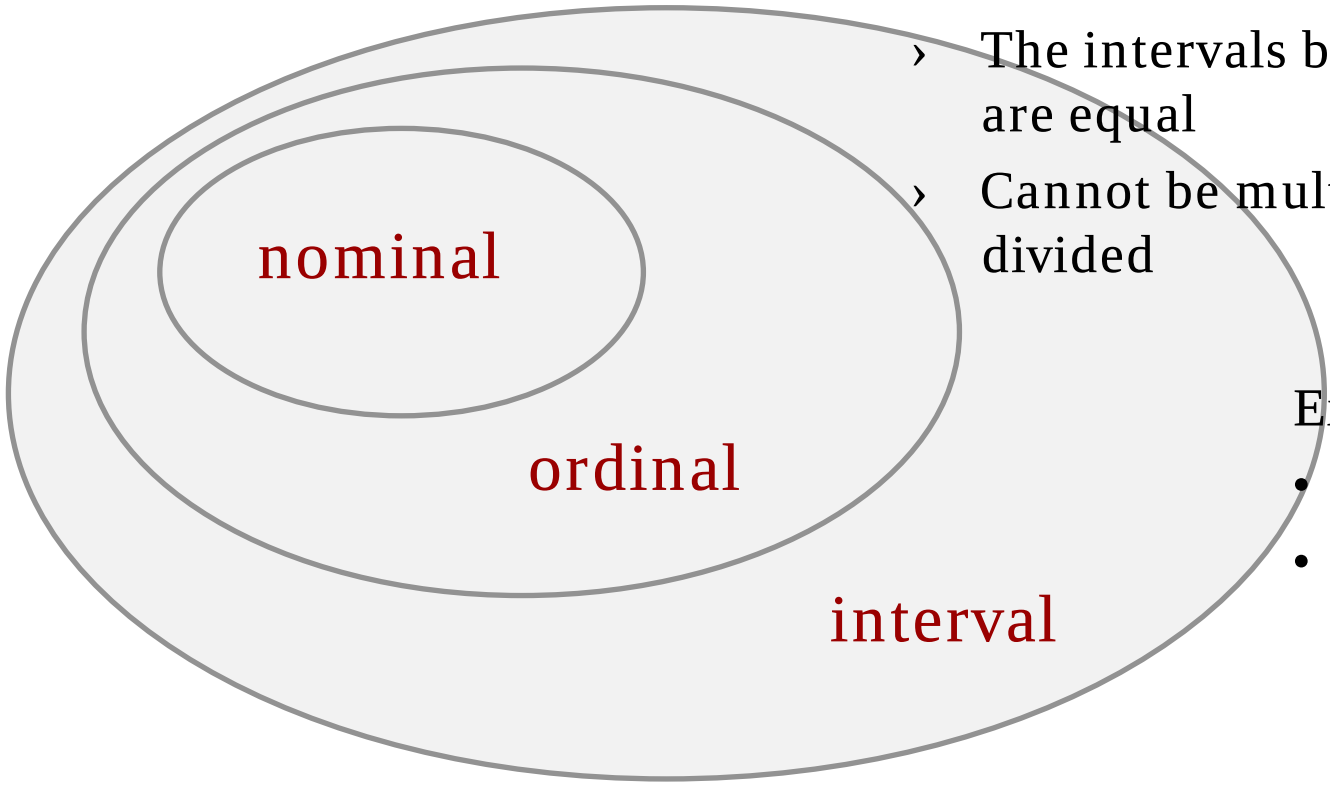
Examples:

- Positions within a company
- Winner, second and third in a race

Types of data



Types of data

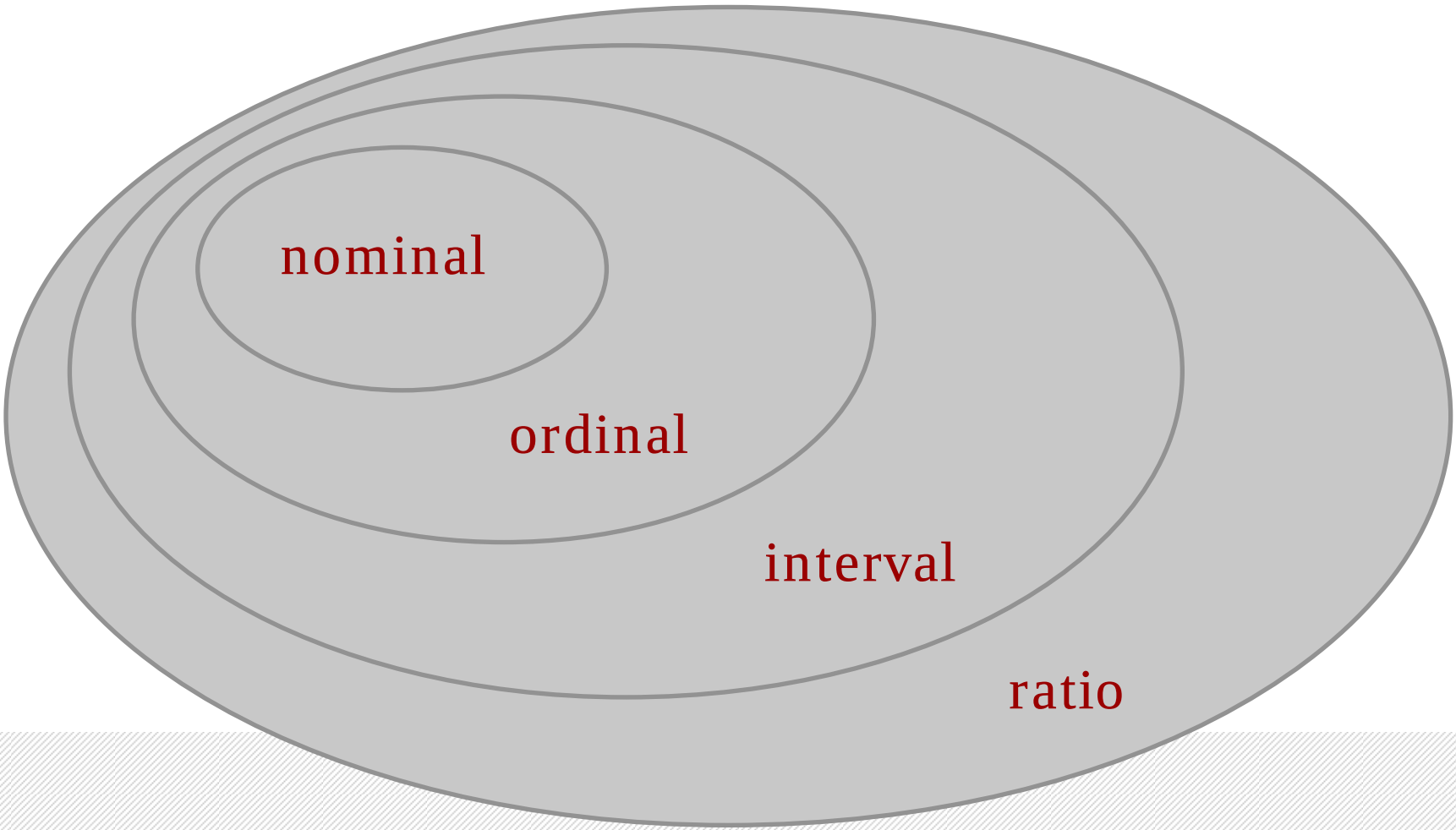


- Interval data: ordered, constant scale, but no natural zero
- The intervals between values are equal
- Cannot be multiplied or divided

Examples:

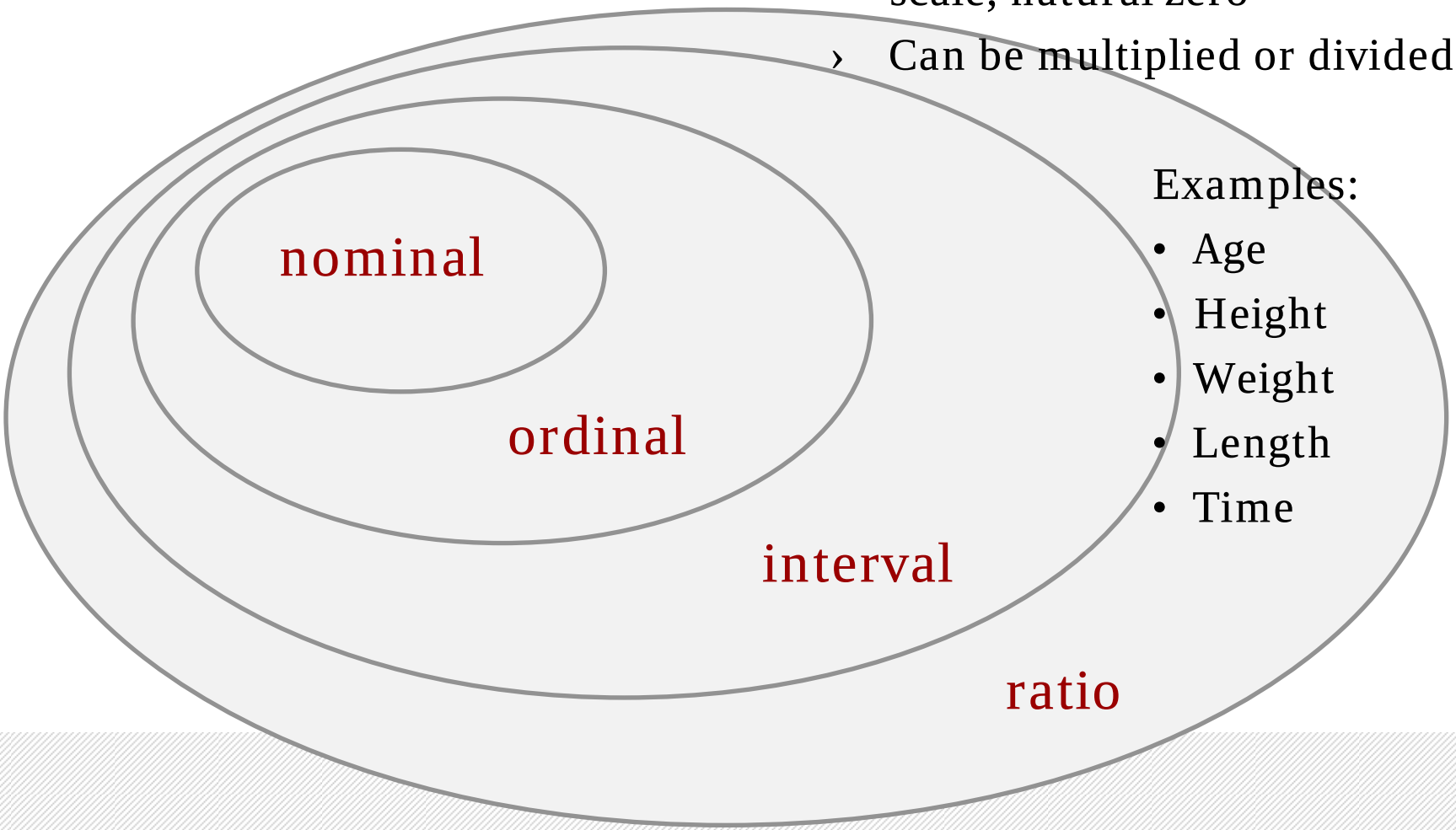
- Temperature
- Dates

Types of data



Types of data

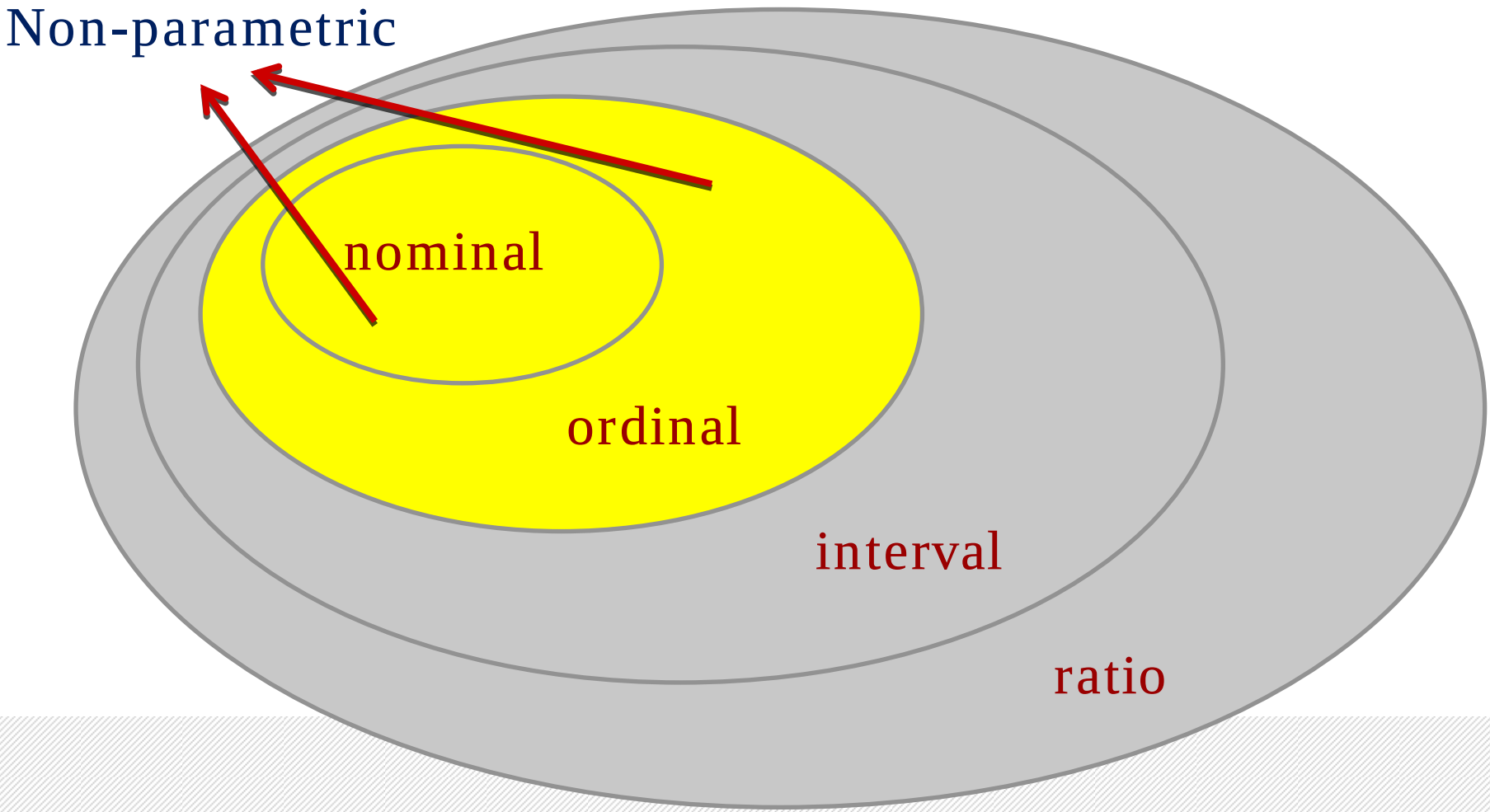
- Ratio data: ordered, constant scale, natural zero
- Can be multiplied or divided



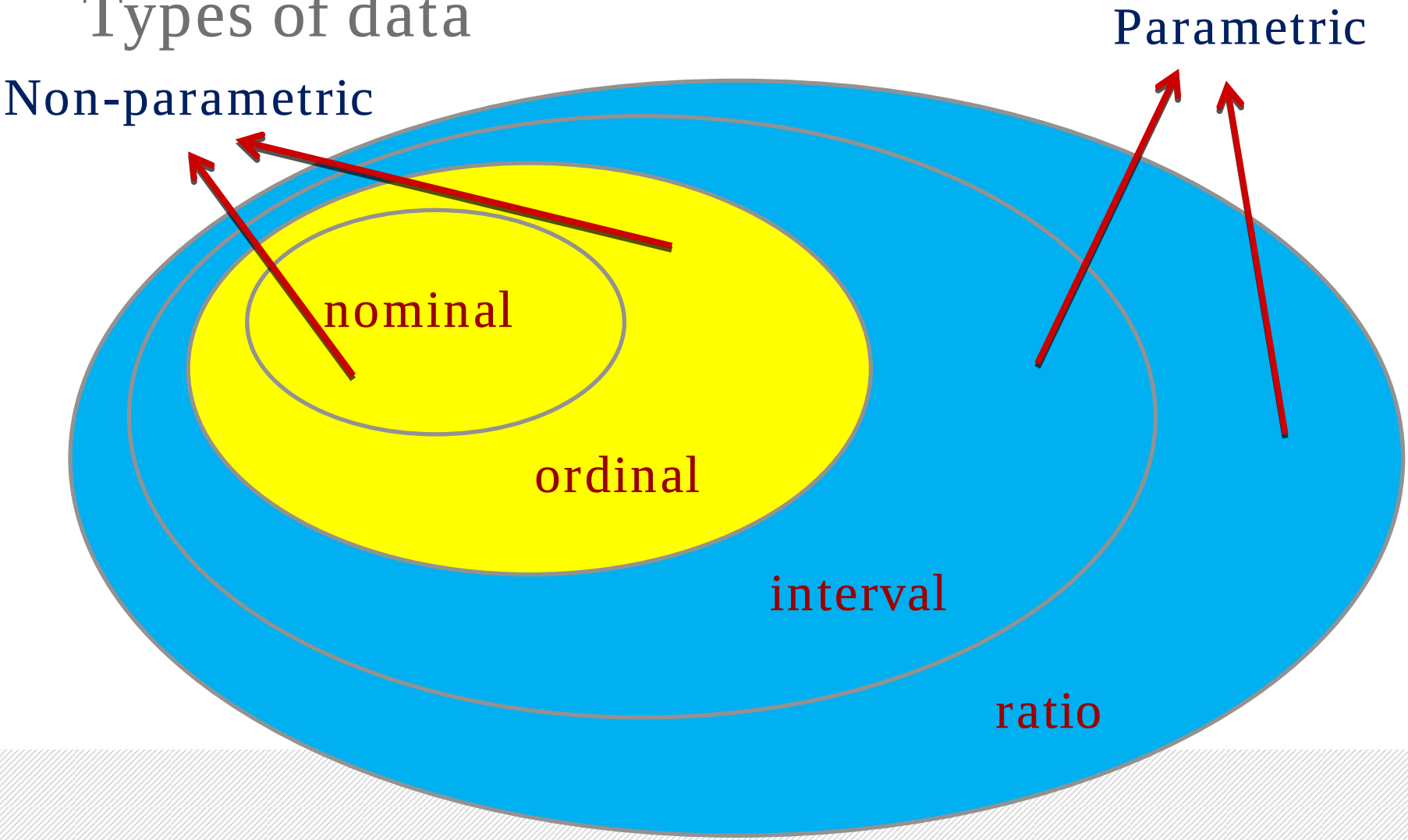
Examples:

- Age
- Height
- Weight
- Length
- Time

Types of data

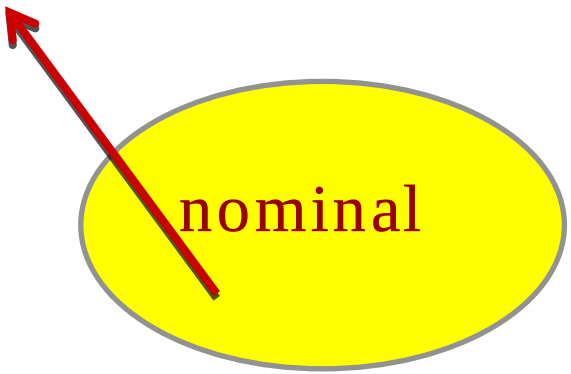


Types of data

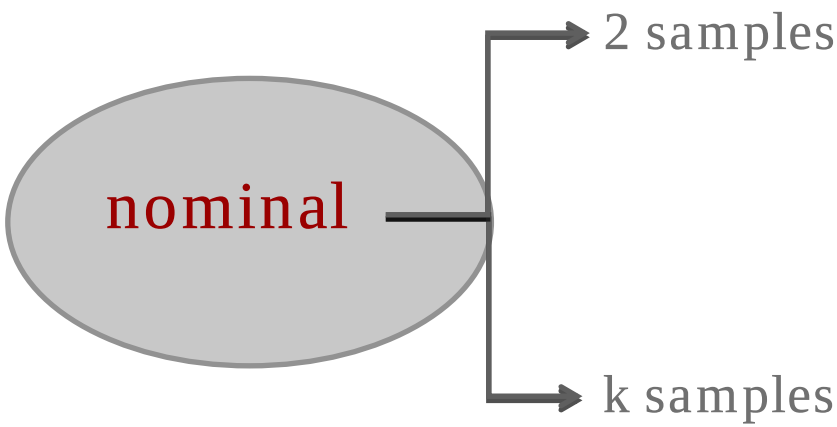


Types of data

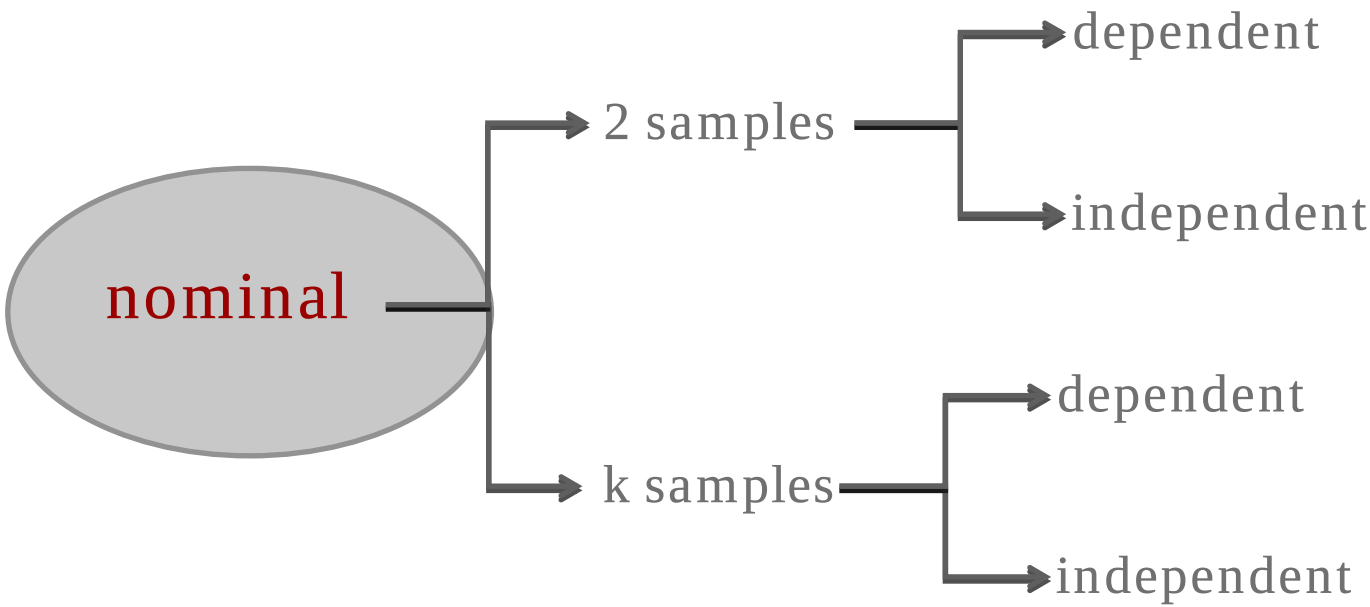
Non-parametric



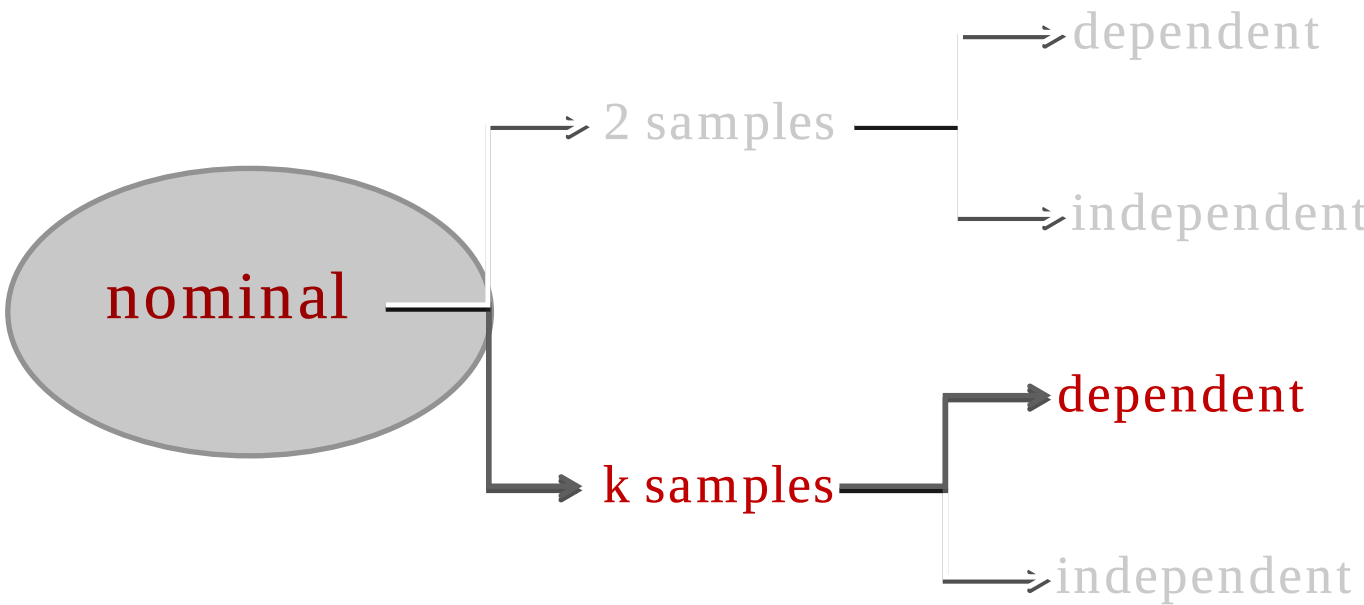
Types of data



Types of data



Cochran's Q test



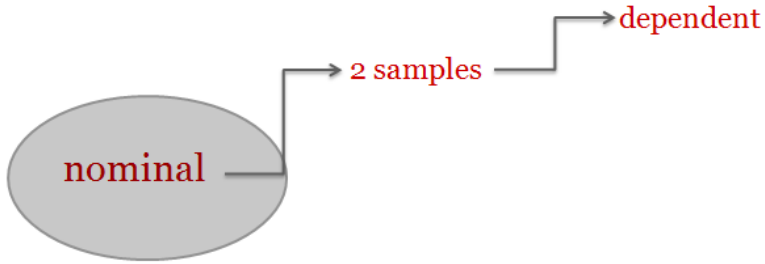
Cochran's Q test: who?

William Gemmell Cochran
(1909 – 1980)

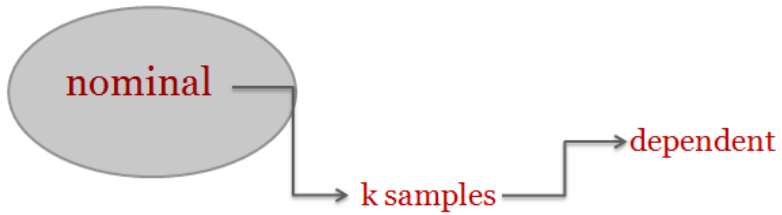
Cochran, W.G. (1950). The Comparison of Percentages in Matched Samples. *Biometrika*, *37*, 256-66.

Cochran's Q test: what?

› McNemar (1947) considered the problem of testing the significance of the difference between two correlated sample proportions



› Cochran suggested a generalization of the problem in which there are *k*(>2) matched samples



Cochran's Q test: what?

- To statistically analyze success rate data

- Tests the hypothesis that several related dichotomous variables have the same mean
- The variables are measured on the same individual or on matched individuals

Cochran's Q test: when?

- Nominal data
- More than 2 samples
- Dependent

- Binary response: succes (1) versus failure (0)

Cochran's Q test: how?

› Each of *k* treatments is independently applied to *b* blocks (or subjects) and each outcome is measured as a success (1) or as a failure (0)

	Treatment 1	Treatment 2	· · ·	Treatment *k*
Block 1	X_{11}	X_{12}	· · ·	X_{1k}
Block 2	X_{21}	X_{22}	· · ·	X_{2k}
Block 3	X_{31}	X_{32}	· · ·	X_{3k}
⋮	⋮	⋮	⋱	⋮
Block *b*	X_{b1}	X_{b2}	· · ·	X_{bk}

Cochran's Q test: how?

- Hypotheses:
  - H_0: treatments are similarly effective
  - H_1: treatments differ in effectiveness

Cochran's Q test: how?

Test statistic: $$T = k\,(k-1) \sum_{j=1}^{k} \left(X_{\bullet j} - \frac{N}{k} \right)^2 \Big/ \sum_{i=1}^{b} X_{i\bullet}\,(k - X_{i\bullet})$$

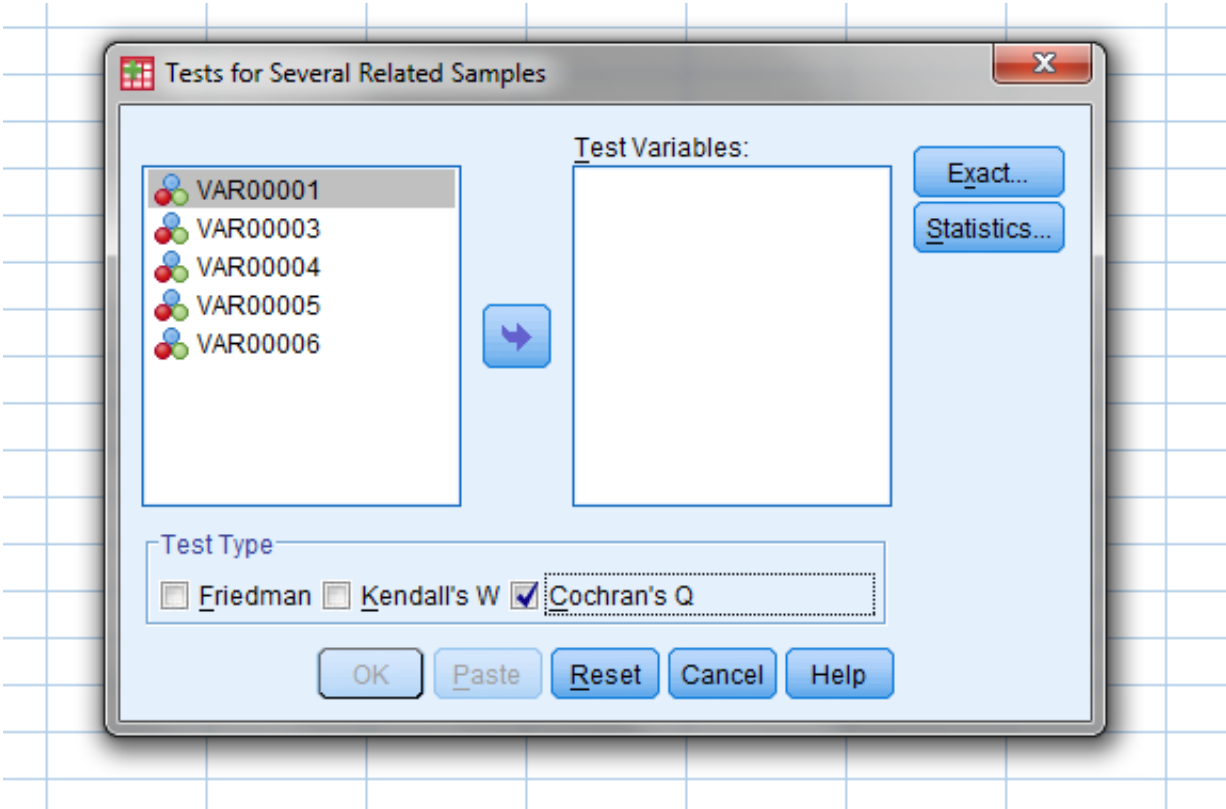
- k is the number of treatments
- $X_{\bullet j}$ is the column total for the j^{th} treatment
- b is the number of blocks
- $X_{i\bullet}$ is the row total for the i^{th} block
- N is the grand total

› For significance level α, the critical region is $T > \mathrm{X}^2_{1-\alpha,k-1}$

where $\mathrm{X}^2_{1-\alpha,k-1}$ is the $(1-\alpha)$-quantile of the chi-square distribution with $k-1$ degrees of freedom

Cochran's Q test: how?

› OR...



Cochran's Q test: application?

- Cochran's Q is often used for meta-analyses, e.g:
  - Is there a difference in treatments (to test)...?
  - Is there a difference in tasks (to test)...?
  - Is there a difference in materials (to test)...?

- But also: are ... used in the same rate or is there a difference?
  - Methods
  - Materials
  - Devices

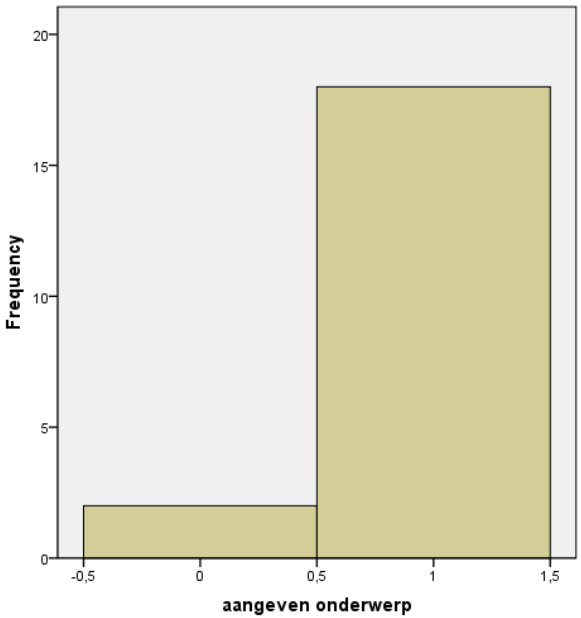
Cochran's Q test: example

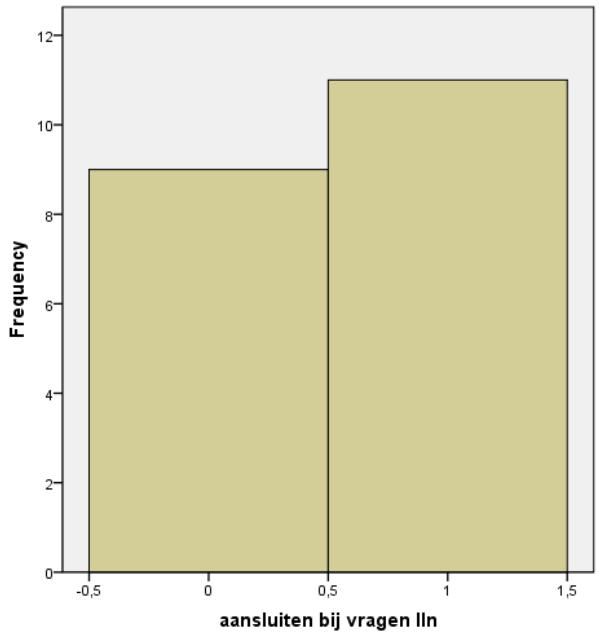
- Background:
  - A teacher has been observed on 20 different moments in his lessons (when he was about to explain something)
  - Observed on 3 different teaching strategies:
    - Point out the subject
    - Correspond with questions of students
    - Use of a stimulating beginning
  - Question: does the teacher use these strategies in the same rate or is there a difference?

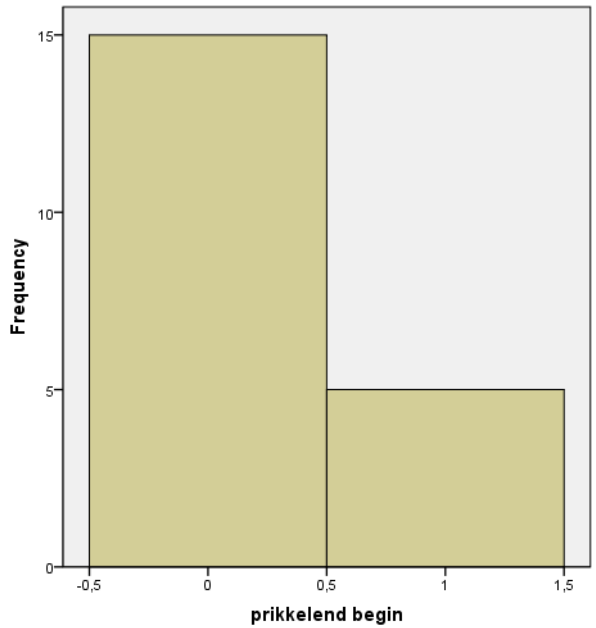
Cochran's Q test: example

	aangevenonderwerp	aansluitenbijvragenlln	prikkelendbegin
1	1	0	0
2	1	0	0
3	1	1	0
4	1	0	1
5	1	1	0
6	0	1	0
7	1	0	0
8	1	1	0
9	1	1	0
10	1	1	0
11	1	0	0
12	0	1	0
13	1	0	1
14	1	0	1
15	1	1	0
16	1	0	1
17	1	0	0
18	1	1	0
19	1	1	0
20	1	1	1

Cochran's Q test: example







Cochran's Q test: example

Frequencies

	Value	
	0	1
aangeven onderwerp	2	18
aansluiten bij vragen lln	9	11
prikkelend begin	15	5

Test Statistics

N	20
Cochran's Q	13,368[a]
df	2
Asymp. Sig.	,001

a. 1 is treated as a success.

- There is a significant difference in the use of teaching strategies used by this teacher ($X_2(2) = 13.37$, $p = .001$)

Cochran's Q test: example

- When you find any significant effect, you need to do a post-hoc test (as you do for ANOVA)

- For Cochran's Q test: run multiple McNemar's tests and adjust the p values with Bonferroni correction *(a method used to address the problem of multiple comparisons, overcorrects for Type I error)*

Cochran's Q test: example

aangeven onderwerp & aansluiten bij vragen lln

aangeven onderwerp	aansluiten bij vrag...	
	0	1
0	0	2
1	9	9

aangeven onderwerp & prikkelend begin

aangeven onderwerp	prikkelend begin	
	0	1
0	2	0
1	13	5

aansluiten bij vragen lln & prikkelend begin

aansluiten bij vragen lln	prikkelend begin	
	0	1
0	5	4
1	10	1

Test Statistics[b]

	aangeven onderwerp & aansluiten bij vragen lln	aangeven onderwerp & prikkelend begin	aansluiten bij vragen lln & prikkelend begin
N	20	20	20
Exact Sig. (2-tailed)	,065[a]	,000[a]	,180[a]

a. Binomial distribution used.

b. McNemar Test

Cochran's Q test: example

- Bonferroni correction:
  - α = 0.05
  - 3 comparisons
  - 0.05/3= 0,016666666666667

Test Statistics[b]

	aangeven onderwerp & aansluiten bij vragen IIn	aangeven onderwerp & prikkelend begin	aansluiten bij vragen IIn & prikkelend begin
N	20	20	20
Exact Sig. (2-tailed)	,065[a]	,000[a]	,180[a]

a. Binomial distribution used.

b. McNemar Test

- Only the difference between the use of 'point out the subject' and 'stimulating beginning' is significant

Limitations of the Cochran's Q test

- Only determines the occurrence of a change, but does not evaluates the extent of change
  - Possible to do multiple McNemar's tests, but no interaction effect can be measured

Limitations of the Cochran's Q test

- Only determines the occurrence of a change, but does not evaluates the extent of change
  - Possible to do multiple McNemar's tests, but no interaction effect can be measured
- The test is known to be poor at detecting true heterogeneity among studies as significant
  - Meta-analyses often include small numbers of studies, and the power of the test in such circumstances is low
  - Because the test is poor at detecting true heterogeneity, a non-significant result cannot be taken as evidence of homogeneity

Limitations of the Cochran's Q test

- Only determines the occurrence of a change, but does not evaluates the extent of change
  - Possible to do multiple McNemar's tests, but no interaction effect can be measured
- The test is known to be poor at detecting true heterogeneity among studies as significant
  - Meta-analyses often include small numbers of studies, and the power of the test in such circumstances is low
  - Because the test is poor at detecting true heterogeneity, a non-significant result cannot be taken as evidence of homogeneity
- The test does not accommodate a control group, because it is a test for use with dependent observations

Discussion

› To begin with:

Questions?

Discussion

› Discussion points?

› ...for example:

- Is this analysis usefull in Linguistic analyses?
- Is it applicable in your research (e.g. pilot study)?

References

› Cochran, W.G. (1950). The Comparison of Percentages in Matched Samples. *Biometrika*, *37*, 256-66.

› Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*, 557-560.

› Pett, M.A. (1997). *Nonparametric Statistics for Health Care Research: Statistics for Small Samples and Unusual Distributions.* Thousand Oaks, CA: SAGE Publications.