Causal analysis

MEVIT4800 25 March 2009

Tanja Storsul

Causal analysis in practice

- How powerful is media?
  - Media violence affects real-life violence?
  - Media sets the agenda?

Goal of lecture

- Knowledge about basic statistical methods and how they can be used.
- Knowledge about interpretation of causal relations.

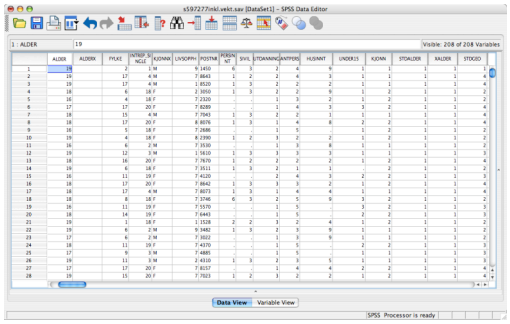
Basic concepts

- **Units of analysis** (enheter)
  - Objects of investigation (can be persons, actions, meanings).
- **Variables**
  - Characteristics of the units.
- **Values (categories)**
  - Variations in the characteristics.

What is what?

"Youth use the Internet for democratic participation to a larger degree than older people."

The data matrix



Data matrix

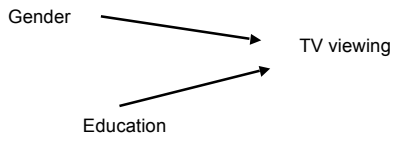
- One purpose of compiling such a data matrix is to study the relationships between the variables, whether a certain value on one variable tends to be combined with specific values on others.
- In quantitative analysis this question is answered by counting how often the various combinations of values on the variables occur in the data matrix.

Hellevik 1988: 2

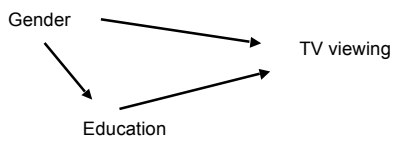
Bivariate associations

	Men	Women	Diff		Low edu	High edu	Diff
High tv cons	(270) 42%	(1120) 74%	-32		(1200) 76%	(190) 32%	44
Low tv cons	(370) 58%	(400) 26%	32		(370) 24%	(400) 68%	-44
	(640) 100%	(1520) 100%			(1570) 100%	(590) 100%	

Causal model



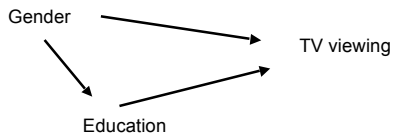
Causal model



Causal concepts

- **Gross association (GA)**
  - Bivariate association between two variables (bivariat sammenheng).
- **Causal effect (CE)**
  - Net association between independent and dependent variable controlled for prior variables in the causal model.
- **Direct effect (DE)**
  - Net association between two variables controlled for all other variables influencing the dependent variable in the causal model.
- **Indirect effect (IE)**
  - Component of the association between two variables which is due to intervening variables in the causal model (CE-DE).
- **Spurious effect**
  - Component of the association between two variables which is due to prior variables in the causal model (GA-KE).

Causal model

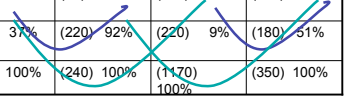


Crosstabulations

	Men		Women	
	Low edu	High edu	Low edu	High edu
High tv consump	(250) 63%	(20) 8%	(950) 81%	(170) 49%
Low tv consump	(150) 37%	(220) 92%	(220) 9%	(180) 51%
Total	(400) 100%	(240) 100%	(1170) 100%	(350) 100%

Differences in proportions

	Men		Women	
	Low edu	High edu	Low edu	High edu
High tv consump	(250) 63%	(20) 8%	(950) 81%	(170) 49%
Low tv consump	(150) 37%	(220) 92%	(220) 9%	(180) 51%
Total	(400) 100%	(240) 100%	(1170) 100%	(350) 100%



Partial associations for gender: Low edu: -18, High edu: -41

Partial associations for education: Men: 55, Women: 32

Statistical interaction

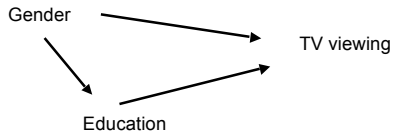
Net associations as weighted average

	Gender	Education
Partial associations	63 - 81 = -18 8 - 49 = - 41	63 - 8 = 55 81 - 49 = 32
Weight (share of all units)	(400 + 1170)/2160 = 0,73 (240 + 350)/2160 = 0,27	(400 + 240)/2160 = 0,3 (1170 + 350)/2160 = 0,7
Net association as weighted average	(-18) * 0,73 = -13 (-41)* 0.27 = -11 - 24	55 * 0,3 = 17 32 * 0,7 = 22 39

Causal analysis

	Gender	Education
Gross association (GA) ♣ (bivariate relationship)	-32	44
Causal effect (CE) (association net of prior variables)	-32	39
Direct effect (DE) ♣ (association net of all causal variables)	-24	39
Indirect effect (IE = CE - DE) (association due to intervening variables)	8	0
Spurious effect (SE = GA - CE) (association due to prior variables)	0	5

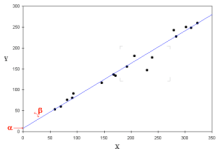
Causal model



Other statistical means

Regression

- A regression coefficient expresses the expected difference in mean value on the dependent variable for units which are one unit of measurement apart on an independent variable
- For example: the expected difference in mean time used of tv for persons who are one year apart in age.



Other statistical methods

Factor analysis:

- Analyses correlations between all variables and shows patterns of correlations or underlying dimensions in the data. Two variables with strong correlations are usually part of the same dimension.
- Often used in research on opinion formation.