SEM 2: Structural Equation Modeling

Week 2 - Causal modeling

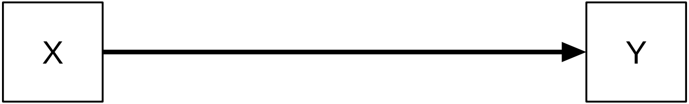
Sacha Epskamp

May 1, 2020

Causal modeling

- This course will introduce structural equation modeling (SEM)
- In SEM, we will discuss modeling complex causal hypotheses
- Again, all variables are assumed normally distributed and all associations are assumed linear
- Causal hypotheses can be specified between observed and latent variables
- CFA is a special case of SEM

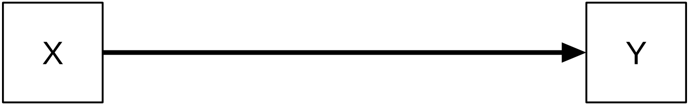
Causal models



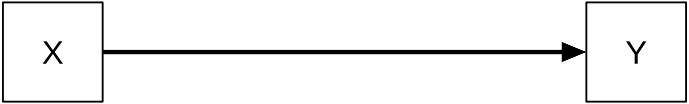
X causes Y

Endogenous and exogenous

- **Exogenous** (independent) variables are variables of which the causal origin are not modeled
  - Exogenous variables have a variance (sometimes not drawn)
  - Exogenous variables, except residuals, are allowed to covary (sometimes not drawn)
  - Latents: ξ (xi); observed: x (x is also used for indicators of latent exogenous variables)
  - Residuals are exogenous
- **Endogenous** (dependent) variables are variables of which the causal origin are modeled
  - Simply stated: endogenous variables have incoming arrows
  - Endogenous variables do not have a variance by themselves
  - Latents: η (eta); observed: y
  - The causal equation for endogenous variables can be derived from the path diagram by summing all incoming edges

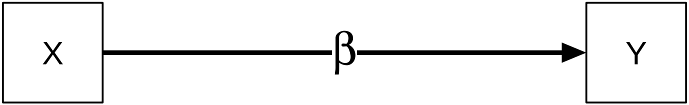


X is exogenous, Y is endogenous

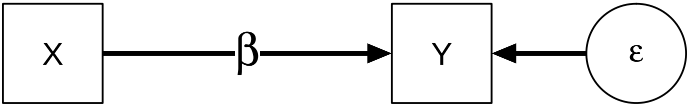


$$y_i = x_i$$

Causal effect goes from right hand side to left hand side. Experimentally changing x will change y, experimentally changing y will not change x

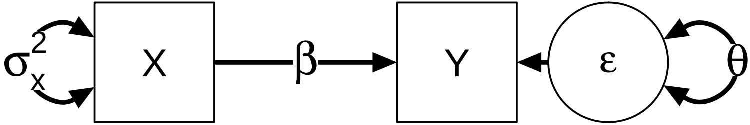


$$y_i = \beta x_i$$

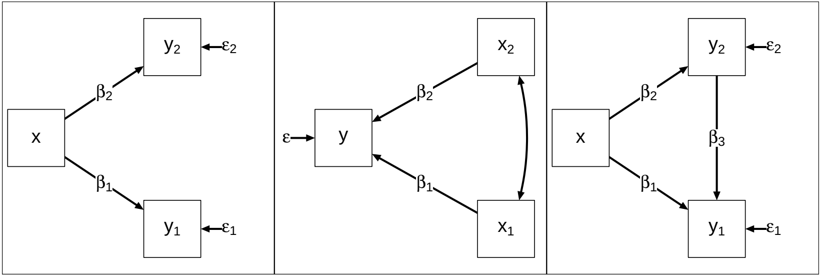


$$y_i = \beta x_i + \varepsilon_i$$

Exogenous variables have a variance (often not drawn)



$$y_i = \beta x_i + \varepsilon_i$$
$$x \sim N(\mu_x, \sigma_x)$$
$$\varepsilon \sim N(0, \theta)$$



Goal: derive a model to explain (co)variances related to observed endogenous variables (including covariances between observed endogenous and exogenous variables), using parameters and (co)variances of observed and unobserved exogenous variables.