Structural Equation Models

Mark Andrews
Psychology Department, Nottingham Trent University

mark.andrews@ntu.ac.uk

What are Structural Equation Models?

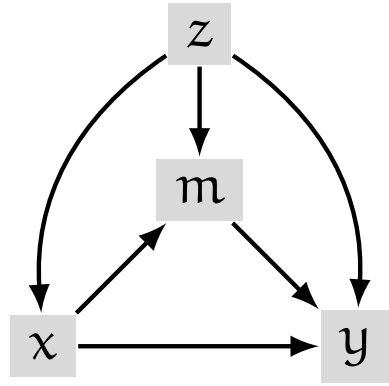
- Structural equation models (SEM) are
  1. Systems of statistical regression models . . .
  2. . . . that represent hypothetical causal models.
- As regression models, they generally include *latent* variables.
- Path analysis and SEM can be seen as either identical, or SEM can be seen as path analysis with additional latent variables.

A very brief history

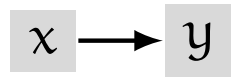
- SEM began with the path analysis work of biologist Sewall Wright around 1920.
- SEM became more widely used from 1970s onwards, but the focus became centred on algorithms and computer implementation, the peculiarities or linear-normal models, and the causal interpretation of SEM was downplayed or eschewed.
- Since 2000s, SEM has become increasingly associated with *causal graph theory* (see J. Pearl), and no longer focused on linear-normal models.

SEM & DAG

- A SEM model can be represented by a *directed acyclic graph* (DAG), where vertices represent variables and directed edges represent causal relationships
- For example,



- For example, we interpret the following to mean that x causes y:



Regression models and DAGs

- Any regression model, or probabilistic model, can be represented by a DAG.
- All variables, including parameters, are represented by vertices and directed edges represent conditional statistical dependence.
- A sem is often represented by a *nonparametric* DAG, i.e., where the parameters and functional form of the regression model are not represented.
- Crucially, directed edges represent causal relationships.

Overview of course

- Regression models
- Mediation models, including causal mediation
- Latent variable models
- SEM with `lavaan` and `blavaan`
- SEM with `stan`, `brms`, etc.