

Path Analysis: A Crash Course

REN R 690 – April 4, 2018
 Melanie de Kappelle and Jessica Grenke

Overview

- What is path analysis?
 - Terms
 - Introduction to the step-method
- Process
 1. Specification
 2. Identification
 3. Estimation
 4. Evaluation
 5. Re-specification
 6. Interpretation

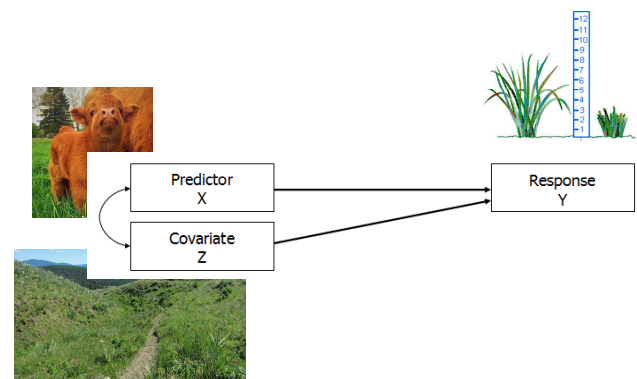
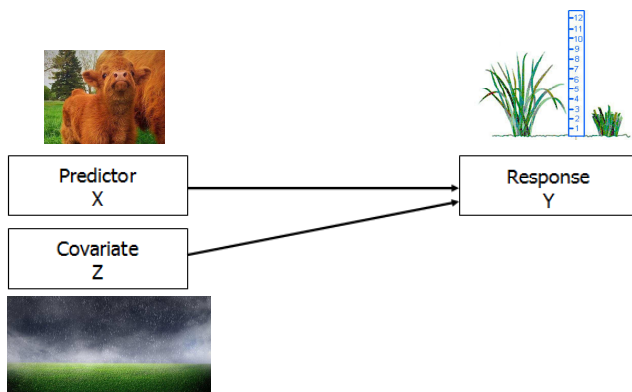


Structural equation modelling (SEM)

- Umbrella term!
- Methods of exploring the “chain of events” between predictor and outcome

Can include observed and non-observable variables

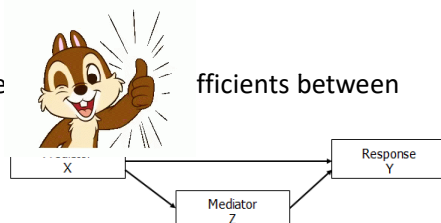
GG!



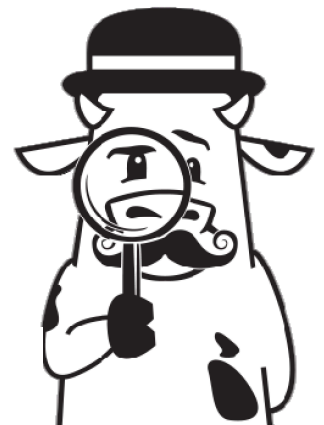
Path Analysis

Structural equation modelling using only observable variables

Based on standardized factors
 = **path coefficient**



Where to begin?



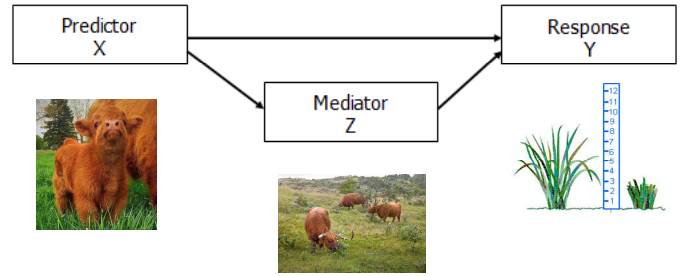
Step 1: Specification

Formally define your hypothesized relationships

- Define variables
- Define relationships between all variables

→ Strength of inference directly dependent on strength of confirmatory nature of analysis

Step 1: Specification



Step 2: Identification

Check that your dataset is sufficient to answer your hypothesized network.

Rules of thumb:

- 1) Residuals are normally distributed
- 2) More observations than variables
- 3) Directional nature constant within each variable
- 4) No error in measurement
- 5) No loops between variables

Step 3: Estimation

Run calculations and analysis
i.e. lavaan package

```

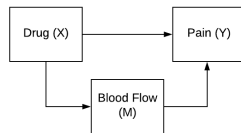
model <- 'Y ~ c*X #direct effect: DRUG has direct effect on Pain#
M ~ a*X #mediator DRUG has on effect on FLOW #
Y ~ b*M #mediator: FLOW has an effect on PAIN#
indirect := a*b #indirect effect#
direct :=c
total:= c +(a*b)'#total effect = sum of direct and indirect#
  
```

Step 4: Evaluation

Gauge degree of agreement between hypothesized relationship and actual results

Parameter Estimates:

Information	Estimate	Std. Err.	z-value	P(> z)	Std. lv	Std. all
Regressions:						
Y ~	(c) 0.092	0.105	0.873	0.383	0.092	0.074
X ~	(a) 0.450	0.109	4.145	0.000	0.450	0.383
M ~	(b) 0.629	0.089	7.048	0.000	0.629	0.595
Variations:						
.Y	Estimate 1.083	Std.Err 0.153	z-value 7.071	P(> z) 0.000	Std.lv 1.083	Std.all 0.607
.M	Estimate 1.358	Std.Err 0.192	z-value 7.071	P(> z) 0.000	Std.lv 1.358	Std.all 0.853
R-Square:						
Y	Estimate 0.393					
M	Estimate 0.147					
Defined Parameters:						
indirect	Estimate 0.283	Std.Err 0.079	z-value 3.573	P(> z) 0.000	Std.lv 0.283	Std.all 0.228
direct	Estimate 0.092	Std.Err 0.105	z-value 0.873	P(> z) 0.383	Std.lv 0.092	Std.all 0.074
total	Estimate 0.375	Std.Err 0.119	z-value 3.160	P(> z) 0.002	Std.lv 0.375	Std.all 0.301

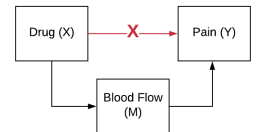


Step 5: Re-specification

If results are a poor fit, go back to Step 1 and adjust hypothesis relationship

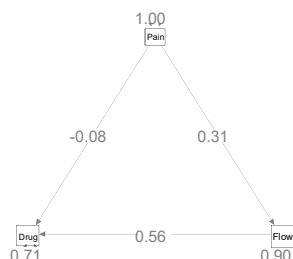
Parameter Estimates:

Information	Estimate	Std. Err.	z-value	P(> z)	Std. lv	Std. all
Regressions:						
Y ~	(c) 0.092	0.105	0.873	0.383	0.092	0.074
M ~	(a) 0.450	0.109	4.145	0.000	0.450	0.383
Y ~	(b) 0.629	0.089	7.048	0.000	0.629	0.595



Step 6: Visualization and Interpretation

- Graphically present results
- Infer relationships
- semPlot package



Questions?

