Tour of Major Correlational Models/Questions

- Still the first & most important is picking the correct model…
- Simple correlation questions
  - simple correlation
  - comparing a correlation across populations
  - comparing correlated correlations
- Multiple regression questions
  - multiple correlation
  - comparing nested multiple regression models
  - comparing non-nested multiple regression models
  - comparing multiple regression models across populations
  - comparing multiple regression models across criterion variables
- Path and Mediation Models

Picking the right statistical model…

Started out simple … What kinds of variables ???

2 quant → correlation 2 qual → X^2 1 @ → ANOVA

Got more interesting with larger ANOVA designs…

- What kinds of variables ?
- What kind of design – How many IVs? BG, MG or WG ?
- What kind of RH: -- main, simple or interaction effect?

Now .. Different models, different "keys," but same importance:

- the analysis must give the most direct test of RH: possible…
- if you pick the wrong model the analyses are worthless!!!

Survey of Major Correlational Models/Questions

- simple correlation questions
  - obtaining and comparing bivariate correlations
- statistical control questions
  - what "would be" the bivariate correlation is all participants had the same score on some "control variable"?
- multiple correlation questions
  - obtaining and comparing models with multiple predictors
- path analysis & mediation
  - Temporal/causal flow & direct vs. indirect effects

Simple Correlation questions (old friends)

$r_{y,x1}$ simple correlation of y and x1

$r_{y,x1}$ vs. $r_{y,x2}$ comparing "correlated" correlations within a population/group (uses Steiger's Z-test)

$r_{y,x1}$ vs. $r_{y,x1}$ comparing the same bivariate correlation in 2 populations/grps (uses Fisher's Z-test)

Examples…

Is there a relationship between # therapy sessions and symptomatic improvement? $r_{imp,\#ses}$

Is # therapy sessions a better predictor of symptomatic improvement than initial level of depression? $r_{imp,\#ses}$ vs. $r_{imp,init}$

Is # therapy sessions a better predictor of symptomatic improvement for adults than for adolescents? $r_{imp,\#ses}$ for adults vs $r_{imp,\#ses}$ for adolescents.

Which type is each of the following?

Does amount of practice predict performance better for novices that for experiences individuals? — Comparing r across populations

Does amount of practice predict level of performance? — Simple r

Does amount of practice predict performance better then prior experience? — Comparing correlated r's

Multiple Regression questions

$R^2_{y.x1,x2,x3,x4}$ multiple correlation with y as the criterion and x1, x2, x3 and x4 as predictors predictors to right of "."

$R^2_{y.x1,x2,x3,x4}$ vs. $R^2_{y.x1,x2,}$ comparing nested models (uses R^2 change F-test)

$R^2_{y.x1,x2}$ vs. $R^2_{y.x3,x4}$ comparing non-nested models (uses Steiger's Z-test)

$R^2_{y.x1,x2,x3,x4}$ vs. $R^2_{y.x1,x2,x3,x4}$ comparing the same multiple regression model in two different populations (uses Fisher's Z-test & Steiger's Z-test)

$R^2_{y.x1,x2,x3,x4}$ vs. $R^2_{z.x1,x2,x3,x4}$ comparing the same multiple regression model with two different criterion, in the same population (Steiger's Z-test)

Examples…

Symptomatic improvement is predicted from a combination of # sessions, initial depression and age. $R^2_{imp.\#ses,init,age}$

Symptomatic improvement is predicted from a combination of # sessions, initial depression and age and prediction is improved by adding # of prior therapists. $R^2_{imp.\#ses,init,age}$ vs. $R^2_{imp.\#ses,init,age,\#ther}$

Symptomatic improvement is predicted better from a combination of # sessions, initial depression and age than from # sessions & # of prior therapists. $R^2_{imp.\#ses,init,age}$ vs. $R^2_{imp.\#ses,\#ther}$

Symptomatic improvement is predicted from a combination of # sessions, initial depression and age better for adults than for adolescents. $R^2_{imp.\#ses,init,age}$ for adults vs. for adolescents

A combination of # sessions, initial depression and age predicts symptomatic improvement better than it predicts treatment satisfaction. $R^2_{imp.\#ses,init,age}$ vs. $R^2_{tsat.\#ses,init,age}$

Which type is each of the following?
Use the notation & tell the test used for each model comparison

Do practice, prior skill and motivation predict performance? single model $R^2_{perf.prac,skill}$

Do practice, prior skill and motivation predict performance on a speeded task as well as they they predict performance on an accuracy task? single model for 2 criterion H & S $R^2_{speed.prac,skill}$ vs. $R^2_{acc.prac,skill}$

Do practice, prior skill and motivation predict performance as well as do prior skill and motivation? nested model comparisons R^2-Δ F-test $R^2_{perf.prac,skill,mot}$ vs. $R^2_{perf.prior,mot}$

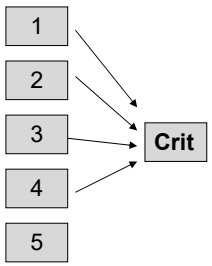
Do practice, prior skill and motivation predict performance as well as do practice, motivation and age? non-nested models H & S $R^2_{perf.prac,skill,mot}$ vs. $R^2_{perf.prac,mot,age}$

Do practice, prior skill and motivation predict performance as well for amateurs as for professionals? R^2 for 2 populations F's Z-test - S $R^2_{perf.prac,skill,mot}$

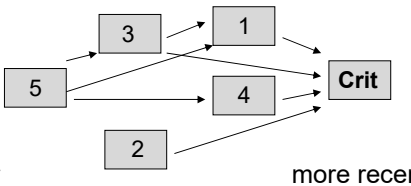
Path Analysis

One way to "think about" path analysis is as an improvement to multiple regression. In addition to asking how the various predictors related to the criterion, path analysis asks how the predictors related to each other on the "temporal/causal path"

"Structure" of a MR model – with hypotheses about which predictors will contribute



A proposed structure for the colinearity among the predictors and how they relate to the criterion – with hypotheses about which paths will contribute

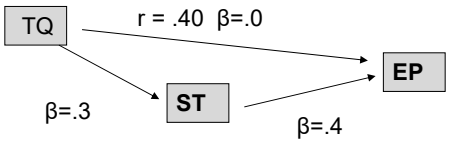


earlier more recent

distal cause proximal cause

Mediation Analysis

Mediation analyses can be thought of as a variety of path analysis that focusses on understanding whether the apparent effect of one variable upon another, based on their temporal precedence pattern and simple correlation, represents a direct effect or one that is mediated by an indirect effect involving another variable.



Here the purpose of the analysis is to consider the r=.4 of Teaching Quality and Exam Performance, when one includes the potential intermediate variable Study Time.

If $\beta_{TQ} = .00$ → complete mediation
If $.00 < \beta_{TQ} < r_{TQ}$ → partial mediation
If $\beta_{TQ} = r_{TQ}$ → no mediation

Path analyses usually refer to examination and hypothesis testing about "direct effects" and "indirect effects".

Mediation analyses usually refer to "intermediate variables".

Examples…

Motivation to succeed has both direct and indirect effects in a model of classroom learning. However, teaching quality does not have direct effects, but does have indirect effects through exam study time.	Path Analysis
Exam study time is an important intermediate variable when considering the relationship between teaching quality and classroom performance.	Mediation Analysis

Which is each of the following?

How good an employee's manager is will be an important intermediate variable when considering the relationship between prior sales experience and sales success.	Mediation Analysis
Prior sales experience will have both direct and indirect effects upon sales success.	Path Analysis
Number of therapeutic sessions Therapeutic engagement will have indirect effects upon therapeutic outcome through therapeutic engagement, but will not have direct effects.	Path Analysis
You have to consider therapeutic engagement as an intermediate variable between number of sessions and therapeutic outcome.	Mediation Analysis