Measurement Theory

Measurement, Scaling, and Dimensional Analysis
2019 ICPSR Summer Program
Prof. Adam M. Enders

Housekeeping

- Course webpage: `www.adamenders.com/teaching`
  - Click "Scaling and Dimensional Analysis"
  - Password: "ICPSR2019"
- Lab on...Friday? 5:30-6:30????
  - Will focus on summated rating model and reliability (tomorrow/Friday)
  - Also: how to work with data in R, and using scales in subsequent analyses

Basics of Measurement

- As long as there have been people, people have been measuring things
- No one really gave this much thought; in fact, not much thought is given by social scientists today
- Throughout history, measurement was a counting operation
- People started giving it thought when they wanted to measure things that were not amenable to a simple counting process
- S.S. Stevens: **measurement is the process of assigning numbers to objects in meaningful ways**

How to Measure Things

- Measurement begins with a classification process; its about categorizing
- When we conduct measurement, we start with a set of objects to be measured and a property (characteristic) of those objects that we can use to distinguish objects from one another
    - So, strictly speaking, we measure characteristics of objects
- How to conduct measurement:
    1. Place objects into mutually exclusive categories; objects that are the same go in the same category and vice versa
    2. Assign numbers to the categories (this is the measurement part)
        - **NOTE:** The numbers assigned to the categories must correspond in some specified way to the differences between the objects with respect to the characteristic that got them categorized in the first place

Measurement as Theory Testing

- The process of measurement involves constructing a **formal model** of our observations
  - By formal model, we mean a subset of the real number system
  - Furthermore, we know we are modeling because measurement, like statistical models, are artificial abstractions (and simplistic ones, at that)
  - Since we are *modeling*, there is going to be an issue with accuracy/error/uncertainty
- **All measurement is theory-testing** (because of the construction and evaluation of models is reliant on theory)
  - So, all measurement is a tentative statement, or conjecture, about the state of reality
  - That is, measurement is falsifiable
  - No such thing as the "true" measure of an object/characteristic (even if it works really well, and over time)

Measurement and Statistical Modeling

- General process of data analysis: Statistical Model $\longrightarrow$ Measured Quantitative Values $\longrightarrow$ Observations
- In other words, we are testing two types of theories when fitting a statistical model
- A poorly fitting model could be due to:
  1. Bad model specification (e.g., unsupported theory, functional form)
  2. Bad measurement – the numerical values are not appropriately assigned to the distinct categories of our variables
- Can help this indeterminacy by being more careful about measurement at outset, or at least taking it seriously as part of the model (more below)

Quantitative vs. Qualitative

- Does all measurement *need* to be quantitative? NO!
  - Example: alphabetization
- No real difference between "qualitative" and "quantitative" data/variables
  - Only one trivial difference: quantitative uses numbers and qualitative doesn't
- Furthermore, there is no such thing as an *empirical* continuous variable
  - If the first step of measurement is categorization, it's impossible to construct a truly continuous variable
  - All *empirical* variables are, in fact, categorical and discrete

Quantitative vs. Qualitative

- In many cases where measurement is not quantitative, a quantitative version could be more useful
  - Consider grades at the university. A+, A, A-, B+, B, B- and so on
  - We can use this qualitative measurement and translate it to quantitative by assigned numerical values to the alphabetic categories. 4.0, 3.5, 3.0, 2.5 and so on
  - What's the difference?
- **Numbers are flexible**
  - We can perform mathematical and statistical operations
  - Can't average C- and A. Can average 3.75 and 2.5.
  - Basic laws and principles of mathematics are firmly established – can use these to make inference easier after measurement has been conducted

Levels of Measurement

- S.S. Stevens invented the modern conception of the levels of measurement
- The "big four" levels of measurement are not the only ones that can exist – there are infinite levels
- LOM refers to differences in how we interpret the numbers we assign to objects
- LOM refers to the nature of the **function that maps from empirical observations to numeric values**

Levels of Measurement

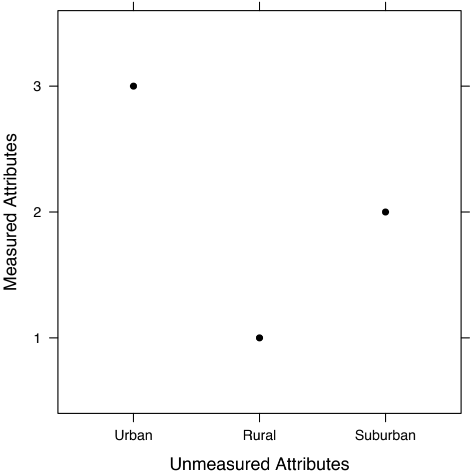
- Differences in functions (note that "$\approx$" means "empirically equivalent"):
  - Nominal: identity-preserving function
    - $A \approx B \implies M(A) = M(B)$
    - $A \not\approx B \implies M(A) \neq M(B)$
  - Ordinal: asymmetry-preserving (monotonic) function
    - $A \approx B \implies M(A) = M(B)$
    - $A \prec B \implies M(A) < M(B)$
  - Interval: specific numeric function
    - $M(A) = c + d(A)$
  - Ratio: same as interval but the intercept $= 0$
    - $M(A) = d(A)$

Levels of Measurement

- Each of the functions that define the 4 basic LOMs can be neatly graphed:
    - Nominal: no real pattern, just preserves identities of observations
    - Ordinal: monotonic curve (graph goes up and right, doesn't change directions)
    - Interval: linear
    - Ratio: linear with intercept at 0

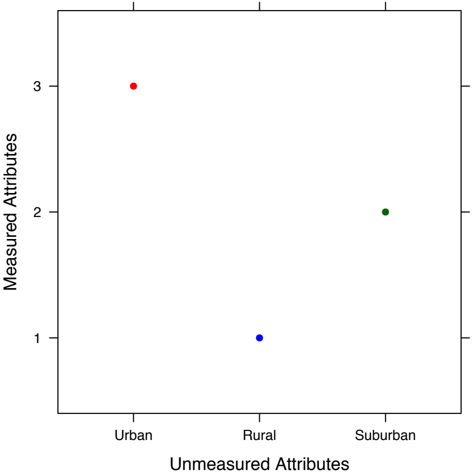
Graphing the LOM: Nominal

No real pattern, just preserves identities of observations



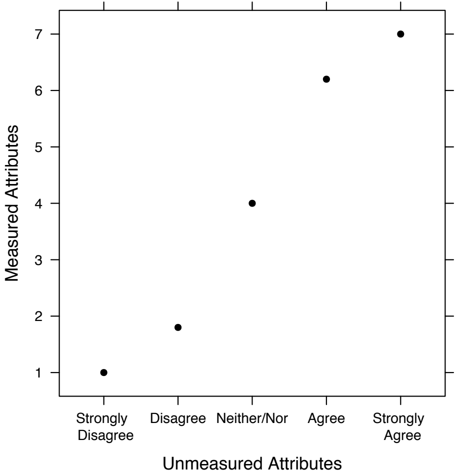
Graphing the LOM: Nominal

No real pattern, just preserves identities of observations



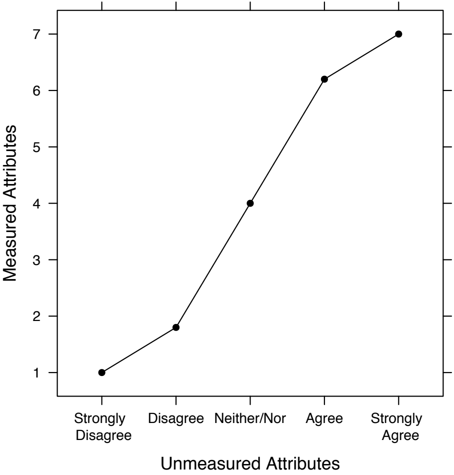
Graphing the LOM: Ordinal

Monotonic curve (graph goes up and right, doesn't change directions)



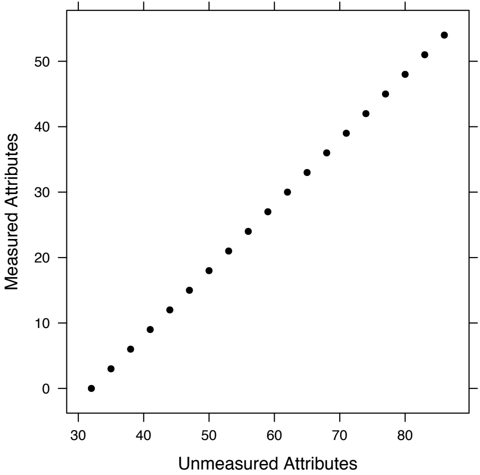
Graphing the LOM: Ordinal

Monotonic curve (graph goes up and right, doesn't change directions)



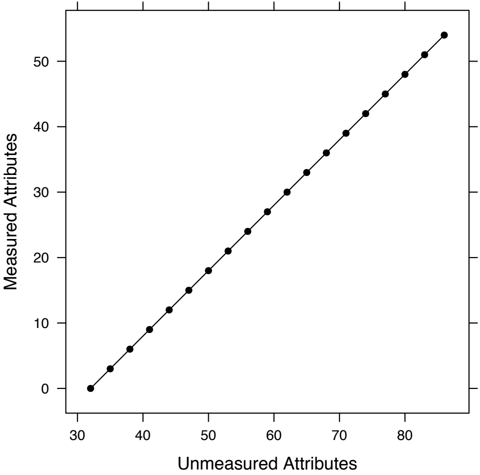
Graphing the LOM: Interval

Linear fit (easiest to understand and most common, but far from only possibility)



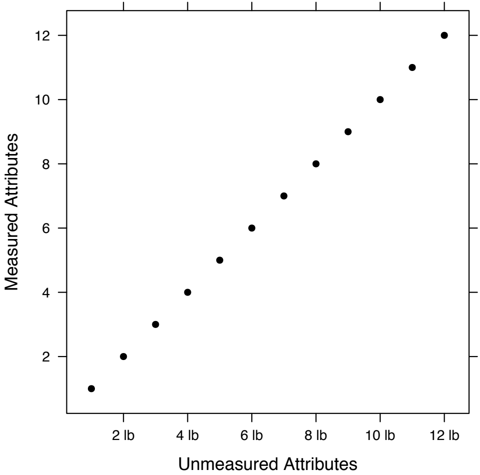
Graphing the LOM: Interval

Linear fit (easiest to understand and most common, but far from only possibility)



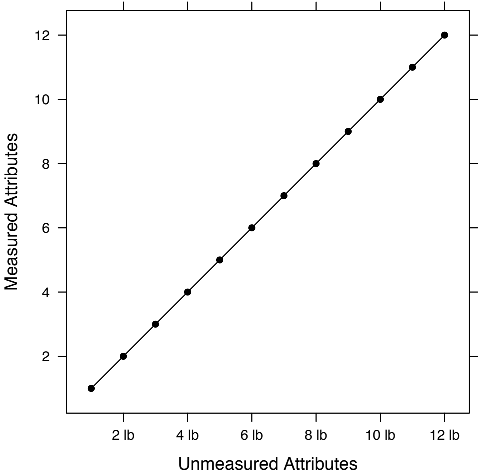
Graphing the LOM: Ratio

Same as interval with 0 intercept



Graphing the LOM: Ratio

Same as interval with 0 intercept



Levels of Measurement

- Measurement *level* only tells us the type of function we're using, *it tells us nothing about the specific set of numeric values we assign to the empirical observations*
  - For example, if you have 5-category ordinal variable you could assign the categories the numbers 0, 10, 425, 6234, 70000 and there would be no problem
- The nature of the function becomes more restrictive as we move up through the levels of measurement
  - There are, for example, fewer values that satisfy the monotonicity required by the ordinal measurement function than there are simply unique identifiers required by the nominal measurement function
  - Parametric functions (e.g., linear) more restrictive than monotonic functions

A Short Exercise in Measurement: Optimal Scaling

- If the numeric values don't really matter much, how do we assign?
  - Convenience! – we usually assign single-digit, successive integer values
- Problem: when we assign convenience scores to variables, we are implicitly subscribing to the assumption that the difference between the numerical values is meaningful
  - Is the "true" difference between Strong Republican and Republican the same as the "true" difference between Weak Republican and Lean Republican?
- We can do better: **optimal scaling**

What is Optimal Scaling?

- Alternative strategy for assigning numbers to objects *after* we've decided the level of measurement, presumably based on theory
- "Optimal scaling is a data analysis technique which assigns numerical values to observation categories in a way which maximizes the relation between the observations and the data analysis model while respecting the measurement character of the data" (Young 1981)
- A basic example:
  - Say you you correlate two variables, a and b, where a is an ordinal variable and b is an interval level variable
  - Why can't you adjust the spacing between categories in a to optimize correlation?

Mechanics of Optimal Scaling

- General procedure:
  1. Estimate the model once with the more restrictive measurement assumptions
  2. Then, re-estimate the model with less restrictive measurement assumptions
     - If the variable behaves the same way under both assumptions, the more restrictive one is appropriate
- Optimal scores have two characteristics:
  1. They have to conform to measurement assumptions we make
  2. Values we assign to the object should maximize fit of the statistical model

Algorithm: ALSOS

- **A**lternating **L**east **S**quares **O**ptimal **S**caling
    1. The variables are assigned initial optimal scale values, and the measurement characteristics are set
    2. Least-squares estimates are obtained for the parameters of the statistical model
    3. If model fit has not improved over the previous iteration, terminate the procedure; otherwise proceed with the following steps
    4. The predicted values from the statistical model are used to generate new optimal scale values for the variables
    5. Return to Step 2 and re-estimate the model using the updated optimally-scaled variable values
- Combine the `opscale` function with iterated OLS regression estimation
- Result: regression model estimates based on variables measured to maximize fit of the model to the data

Implementation in R

- A number of different ways (algorithms) we can "optimally scale" variables
- See documentation for `optiscale` R package for more info, code, and examples
  - Different transformations depending on LOM (e.g., Kruskal's monotonic transformation)
  - Must be used in conjunction with statistical model estimation
- Other R packages:
  - `Gifi`: set of functions to carry out non-linear multivariate data analysis techniques via optimal scaling
  - `aspect`: functions that can be used to maximize some "aspect" of a correlation matrix by optimally scaling original variables
- We will see examples from all of these packages over the course of the semester

Data for Example

- 2012 American National Election Studies
  - High quality, nationally-representative probability sample of U.S. adults
- Variables:
  - `pid`: party identification; 7-point scale ranging from 1 (strong Dem) to 7 (strong Rep)
  - `ideo`: ideological self-identification; 7-point scale ranging from 1 (extremely liberal) to 7 (extremely conservative)
  - `obamaft`: Obama feeling thermometer score; Ranges from 0 (very cold) to 100 (very warm)
  - `romneyft`: Romney feeling thermometer score; Ranges from 0 (very cold) to 100 (very warm)
  - `govpsend`: attitudes about government spending; 7-point scale ranging from 1 (more spending/services) to 7 (less spending/services)
- Theory: ideology, feelings toward major political figures, and attitudes about government spending predict partisanship

Example in R

See "Optimal Scaling Example.r" code

So What?

- Theoretical
  - If we care about the level of measurement, why not seriously integrate it into our models?
  - Measurement is **flexible**
  - Measurement is **theory-testing**, an empirical question
- Practical
  - "If a procedure is known for obtaining a least squares description of numerical (interval or ratio measurement level) data then an ALSOS algorithm can be constructed to obtain a least squares description of qualitative data (having a variety of measurement characteristics)" (Young 1981)
  - Better congruence between measurements and theory, model and data
- Moving forward
  - Can use optimal scaling (the ALSOS routine, in particular) to estimate nonlinear PCA, factor analysis, etc.
  - Will be treating measurement as flexible in our treatment of subsequent scaling models