Quantitative Analysis and Empirical Methods

1) Quantification and Measurement

Jan Rovny

Sciences Po, Paris, CEE / LIEPP

Introduction

- Quantification requires that we measure social phenomena or concepts
- We describe these phenomena using *variables*
- We use various logical techniques of arriving at variables
- We use a set of statistical tools to describe these variables

Quantification

Concepts

- Empirical research analyzes concepts and the relationships between them
- Some concepts are relatively easily measurable
  - 'how many children a woman gave birth to'
  - 'was there a civil war in Paraguay in the 1980s;'
  - 'how did a voter vote in the 2014 European Parliament elections in France'
- Other concepts are much more vague:
  - Academic ability
  - Globalization
  - Liberalism
  - Democracy

Measurement

- Measurement = logically assigning numbers to a phenomenon of interest
- Phenomena are represented by concepts (educational achievement, program success)
- Concepts cannot be measured directly but through observable indicators / variables

Concept

↓

Conceptual definition

↓

Operational definition

↓

Variable

Measurement example

Concept:

Religiosity

↓

Conceptual definition:

"The degree to which an individual adheres to religious teachings and practices"

↓

Operational definition 1:

Survey asking: "To what extent would you say you adhere to religious teachings and practices of a particular religion" (0=not at all ... 10=extremely)

↓

Variable

(0-10)

Measurement example

Concept:
Religiosity
↓
Conceptual definition:
"The degree to which an individual adheres to religious teachings and practices"
↓
Operational definition 2:
Survey asking: "How frequently do you attend religious services" (0=never; 1=only on special holidays/events; 2=few times a year; 3=about once a month; 4=few times a month; 5=weekly; 6=more than once a week)
↓
Variable
(0-6)

Measurement

- There may be different indicators / variables measuring one concept
  - different ways to measure concept (can triangulate)
  - different dimensions of a concept (academic ability: reading, writing, maths...)
- Operational definitions often lack complete congruence with the concept:
  $$\textit{indicator} = \textit{concept} + \textit{error}$$
- Error can be divided into systematic and random error:
  $$\textit{indicator} = \textit{concept} + \textit{error}_{\textit{systematic}} + \textit{error}_{\textit{random}}$$

Components of Observed Measure

$$indicator = concept + error_{systematic} + error_{random}$$

- Concept = the value we are tying to measure
- Systematic error = systematic bias from influences other than the desired concept
- Random error = error from non-systematic influences on the observed measure (noise, buzz)

Example of measurement errors

- Concept = Euroscepticism in France
- Conceptual definition = The extent to which the French oppose European integration.
- Operational definition = Look at latest EP election results and measure the vote share of FN.

Party	%
FN	24.86
UMP	20.80
PS+	13.98

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
- Random error:

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
  - FN vote is anti-establishment vote, i.e. don't like all of them
- Random error:

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
  - FN vote is anti-establishment vote, i.e. don't like all of them
  - FN vote is strategic, i.e. pressure the government
- Random error:

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
  - FN vote is anti-establishment vote, i.e. don't like all of them
  - FN vote is strategic, i.e. pressure the government
  - FN vote is Marine Le Penn vote, i.e. like Le Penn
- Random error:

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
  - FN vote is anti-establishment vote, i.e. don't like all of them
  - FN vote is strategic, i.e. pressure the government
  - FN vote is Marine Le Penn vote, i.e. like Le Penn
  - some Eurosceptics vote for other parties than FN
- Random error:

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
  - FN vote is anti-establishment vote, i.e. don't like all of them
  - FN vote is strategic, i.e. pressure the government
  - FN vote is Marine Le Penn vote, i.e. like Le Penn
  - some Eurosceptics vote for other parties than FN
- Random error:
  - Put the wrong ballot in the envelope

Example of measurement errors

- Systematic error:
  - FN vote is anti-incumbent vote, i.e. don't like the President
  - FN vote is anti-establishment vote, i.e. don't like all of them
  - FN vote is strategic, i.e. pressure the government
  - FN vote is Marine Le Penn vote, i.e. like Le Penn
  - some Eurosceptics vote for other parties than FN
- Random error:
  - Put the wrong ballot in the envelope
  - Turnout skewed due to random events (e.g. severe weather in pro-EU strongholds)

Validity and Reliability

Validity

A *valid* indicator accurately measure the intended concept.

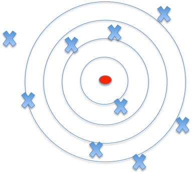
- If a measure is valid, it contains little systematic error.
- Types of validity:
  - Face validity = a measure *a priori* looks like it measures the concept
  - Construct validity = a measure associates strongly with other measures (GRE correlates with GPA)
- Validity cannot be measured directly, but we can triangulate: compare different measures of the same concept (use more data sources, operationalizations etc.)

Reliability

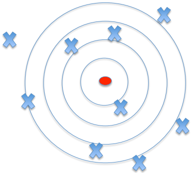
A *reliable* indicator consistently assigns the same value to a constant phenomenon

- If a measure is reliable, it contains little random error.
- In other words, a reliable measure assigns the same value to the same 'thing' every time.
- Reliability can be measured:
    - Test-retest: measure a number of times and compare the results. If they are very similar, then there is little random error, and the measures are reliable.
    - Split-half reliability: take half of (a number of) indicators intended to measure the same concept, and correlate their measure with the other half. High correlation suggests reliability.

Example of Validity and Reliability 1

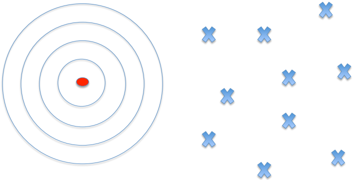


Example of Validity and Reliability 1

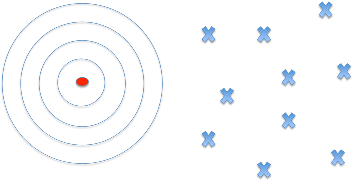


- Valid, unreliable

Example of Validity and Reliability 2

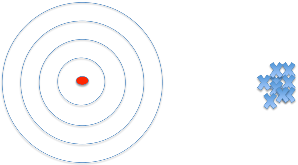


Example of Validity and Reliability 2

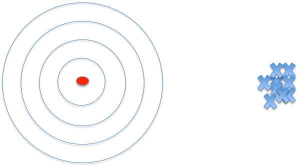


- Invalid, unreliable

Example of Validity and Reliability 3

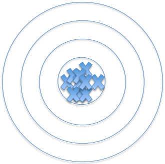


Example of Validity and Reliability 3

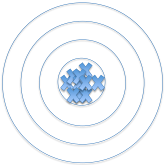


- Invalid, reliable

Example of Validity and Reliability 4



Example of Validity and Reliability 4



- Valid, reliable

Levels of Measurement

Nominal Measures

- Measurement is assigning numeric values to particular attributes according to logical rules.
- There are different types of measurement depending on the phenomenon and rules we use:
- **Nominal Measures**
  - Describe outcomes that fit particular categories.
  - If two objects share the same value of an attribute, they are assigned the same number
  - *male* and *female* are two nominal categories.
  - It does not matter how we assign numeric values: *male=1, female=2* or *female=0, male=1*.

Ordinal Measures

- Describe characteristics of attributes which are logically ordered.
- Can tell that one category is greater or lesser than another
- Cannot tell what the distances between these categories are
- Example 1:
  - Allison likes Prague more than Berlin, but less than Paris.
  - Allison's preferences: *Paris=1, Prague=2, Berlin=3*
  - represents the relationships between Allison's preferences. But *how much more does she like Paris over Prague, than Prague over Berlin*?
- Example 2:
  - Survey question: 'Our country is better off inside the European Union than outside,'
  - Answer options: 'strongly agree', 'agree', 'neutral', 'disagree', or 'strongly disagree': Order but no clear distances!

Interval Measures

- Are based on some underlying continuum
- Allows us to address *how much more* or *how much less* one thing is from another
- **differences in the numbers assigned to attributes reflect the differences between the attributes**
- The numbers assigned to interval measures are arbitrary
- Degrees Celsius: $|-25-0| = |0-25|$
- Degress Fahrenheit shift origin (32 degrees), change scale, but measure same thing
- Transformation of interval measure m: $t(m) = c * m + d$, where c and d are constants.

Scale Measures

- Are like interval measures, but they also have a meaningful zero.
- Income: \$20 per hour v. \$0 per hour.
- Degrees Kelvin ($0 =$ no thermal motion)
- Can transform them only by changing their scales, not origin!
- Transformation m: $t(m) = c * m$, where c is a constant.
- Measure income in cents, rather than dollars. We are multiplying our measure by 100.

Binary Variables

- Also called **dummy variables**
- Take on 0 or 1 only
- Denote two exclusive and exhaustive categories, or the presence or an absence of a phenomenon.
- Every nominal variable can be described by a series of dummy variables
- This plays a particular role in modeling, which you will see later...