Choosing an Approach for a Quantitative Dissertation: Strategies for Various Variable Types

Kuba Glazek, Ph.D.
Methodology Expert
National Center for Academic and Dissertation Excellence
Outline

• Thesis and hypothesis development
• Hypothesis testing techniques
• Types of variables in a study
• Number of variables in a study
• Data analysis techniques
Thesis and Hypothesis

• Thesis: A position or proposition that a person advances and offers to maintain by argument
  – General and conceptual
  – Position/proposition is based on literature review

• Hypothesis: A tentative assumption made in order to draw out and test its logical or empirical consequences
  – Specific and testable
  – Use operational definitions for constructs
  – E.g., “Personality as measured by X is positively correlated with well-being as measured by Y”

http://www.merriam-webster.com/dictionary/
Overview

- Qualitative and quantitative studies’ methods diverge upon determination of thesis

**Qualitative**
- In-depth understanding based on small sample
- No hypothesis *per se* (exploratory)

**Quantitative**
- Broad understanding based on large sample
- Hypothesis-driven (confirmatory)
Null and Alternative Hypothesis Logic

• **Null hypothesis (H\(_0\))**
  – Always the starting point
  – A priori assumption of no effect that has to be overcome, or rejected

• **Alternative hypothesis (H\(_1\) or H\(_a\))**
  – The alternative to a null effect is a *significant* effect
  – If H\(_1\) is supported, then H\(_0\) is rejected
Open-ended vs. Directional Hypotheses

- **Open-ended/two-tailed hypothesis**
  - $H_0$: Null effect
  - $H_1$: Significant association between factor A and factor B (note: direction of effect *not* stated)

- **Directional/one-tailed hypothesis**
  - $H_0$: Null effect
  - $H_1$: Group A significantly higher than Group B (note: direction of effect stated)

- Decision is based on ability to predict results
  - Literature review
  - Identified gaps
Operational Definitions of Support

• Significance testing: $p < .05, p < .001$
  – If, in reality, there is no effect, what is the probability of the obtained effect?
  – If the probability is very low (i.e., less than 5%), then we agree that the effect is real and reject $H_0$

• Effect sizes: Pearson’s $r$, Cohen’s $d$, eta squared ($\eta^2$)
  – Even if not significant, an effect can still be important

• Confidence intervals: 95% CI
  – The expected range of
    • Differences between group
    • Correlation coefficient
    • The significance value
Point Estimates vs. Interval Estimates

http://www.methodspace.com/profiles/blogs/the-joy-of-confidence-intervals
Terminology: Independent and Dependent Variables

- Independent variable (IV)
  - Used in experiments and quasi-experiments
  - Experiment *manipulates* (e.g., type of treatment)
  - Quasi-experiment *compares* (e.g., sex)
  - IV has two or more *levels*
    - Experiment: standard treatment or experimental treatment
    - Quasi-experiment: male or female

- Compare effect of levels of IV on a *dependent* variable (DV)

- For example
  - Number of anxiety attacks in PTSD patients *depends on* treatment type administered to patient
  - Number of anxiety attacks in PTSD patients *depends on* sex of patient
Terminology: Predictors and Outcomes

• Predictor variable
  – Used in **correlational** and **observational** studies
  – Observable phenomena that occur out in the world
    (e.g., strength of attachment, duration of incarceration)

• Predictor used to describe patterns in an
  **outcome** or **criterion** variable

• For example
  – Strength of attachment *predicts* self-esteem
  – Duration of incarceration *predicts* time until recidivism
## Variations on Nomenclature

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Experimental and quasi-experimental design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictor variable</td>
<td>Outcome variable</td>
<td>Non-experimental design</td>
</tr>
<tr>
<td>Antecedent variable</td>
<td>Consequence variable</td>
<td>Multiple types of designs</td>
</tr>
<tr>
<td>Treatment variable</td>
<td>Criterion variable</td>
<td>Multiple types of designs</td>
</tr>
<tr>
<td>Stimulus variable</td>
<td>Response variable</td>
<td>Behavioral psychology</td>
</tr>
<tr>
<td>Cause</td>
<td>Effect</td>
<td>Experimental design</td>
</tr>
<tr>
<td>Bivariate relationship</td>
<td></td>
<td>Correlational design</td>
</tr>
</tbody>
</table>
Theory: Independent and Dependent Variables

• Causality demonstrated by
  – Level of IV is *sufficient* to observe effect on DV
  AND
  – Level of IV is *necessary* to observe effect on DV

• For example
  – Sufficient: *With* experimental treatment, anxiety decreases
  – Necessary: *Without* this treatment, anxiety does *not* decrease
Theory: Predictors and Outcomes

• Correlation does not imply causation

• Causal claims based on correlational models rest on the following assumptions (e.g., Preacher, Rucker, & Hayes, 2007, p. 217)
  – Predictor variable precedes outcome variable in time
  – Correlation between predictor and outcome variable
  – Sources of spurious correlation have been controlled (partial correlations)

• Correlational study needs zero point
  – Demonstrate absence of predicted outcome consistent with absence of predictor

• For example:
  – Pre-treatment data
  – Comparison group
  – Published evidence
Variable Types and Examples

• Categorical or nominal
  – Has qualitatively different categories/levels
  – E.g., veteran status, attachment style

• * Ordinal
  – Has ranked, or ordered categories/levels
  – E.g., SES (lower, middle, upper), education level (elementary, middle, high, college)
  – E.g., Likert scales*

• Interval
  – Has equal increments (i.e., intervals) between categories/levels
  – E.g., Likert scales (strongly disagree – strongly agree)*

• Continuous or ratio
  – Same as interval, but with meaningful zero
  – E.g., Likert scales*, age, number of sessions
Semantic Differentials: Categorical or Continuous?

- Less than five levels should be considered a categorical predictor or outcome
  
<table>
<thead>
<tr>
<th>yes</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td>disagree</td>
<td>neutral</td>
</tr>
</tbody>
</table>

- Five or more *may be* considered a continuous predictor or outcome

- Use seven-point (or more) scale

- Subjectivity problem and non-parametric solution (*Brown-Forsythe test* in place of “regular” ANOVA)
Variables: Flexibility of Data

• Ratio $\rightarrow$ interval $\rightarrow$ ordinal $\rightarrow$ categorical
  – Going in the other direction is not possible
  – E.g., age (ratio) $\rightarrow$ old/young (ordinal)

• Pros
  – Maximize group differences (e.g., oldest 1/4 vs. youngest 1/4)
  – Account for violations of statistical assumptions

• Cons
  – Loss of parametric information
  – 21-25 years (youngest 25% of sample) $\rightarrow$ group 1
  – 28-60 years (oldest 25% of sample) $\rightarrow$ group 2
**Design: Between- and Within-Subjects**

- **Between-subjects design** (cross-sectional studies)
  - Different participants for different levels of IV/predictor
  - E.g., young vs. old; violent offenders vs. non-violent offenders

- **Within-subjects design** (longitudinal studies)
  - Same participants for different levels of IV/predictor
  - E.g., time in treatment; pre- and post-intervention status

- **Mixed design** (combination of both)
  - Different groups at multiple time points
  - E.g., pre- and post-intervention changes in violent and non-violent offenders
Formulation of Hypotheses

• Differences between groups
  – Categorical IV and continuous DV
  – E.g., significantly higher anxiety in females vs. males
  – ANOVAs and t tests

• Degree of association
  – Continuous predictor and continuous outcome; categorical predictor and categorical outcome
  – As number of therapy sessions increases, does anxiety decrease?
  – Correlation indicates the significance and strength of a relationship
  – How much improvement can be expected for each additional session?
  – Linear regression allows for prediction
## Decision Matrix for Quantitative Analyses

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Categorical</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>One</td>
<td>One</td>
</tr>
<tr>
<td>None</td>
<td>Frequency Distribution  $\chi^2$ Goodness of Fit</td>
<td>Descriptive Statistics</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$ Test of Independence</td>
<td>Shape of Distribution</td>
</tr>
<tr>
<td>One</td>
<td>$\chi^2$ Test</td>
<td>2 groups: $t$ Test</td>
</tr>
<tr>
<td></td>
<td>Fisher’s Exact Test</td>
<td>&gt; 2 groups: ANOVA (or regression)</td>
</tr>
<tr>
<td>Multiple</td>
<td>Log-linear Model</td>
<td>ANOVA</td>
</tr>
<tr>
<td></td>
<td>Logistic Regression</td>
<td>ANCOVA (or regression)</td>
</tr>
<tr>
<td>One</td>
<td>Logistic Regression</td>
<td>Correlation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Linear Regression</td>
</tr>
<tr>
<td>Multiple</td>
<td>Logistic Regression</td>
<td>Multiple Regression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Structural Equation Modeling</td>
</tr>
</tbody>
</table>
Cleaning Collected Data

• Enter data and check for integrity
  – Transcription errors are (statistically) **inevitable**
  – Check for missing and erroneous values

• Test statistical assumptions
  – Normality of distributions
  – Homogeneity of variance

• Address violations of assumptions
  – Transform variables
  – Remove influential outliers
  – Use non-parametric tests

http://www.unc.edu/~nielsen/soci709/m3/m16008.gif
Summary

• Start with a testable prediction
  – Can test strength of association between variables
  – Can test effect of one variable on another
• Can use categorical and/or continuous variables
• Appropriate tests for any combination of variables
  – One or multiple predictor variables
  – One or multiple outcome variables
Qualitative Methods Webinar

• Qualitative and quantitative studies’ methods diverge upon determination of thesis

• Philip Adu, Ph.D. will present on qualitative methods tomorrow (2/4/14)
  – 10:00 AM Pacific (1:00 PM Eastern, Noon Central)
  – https://www3.gotomeeting.com/join/261094622
Thank You!

Questions?

• Video will be available on NCADE YouTube channel: https://www.youtube.com/channel/UChCvwPnugSXf85vbQfSeLJQ

• NCADE website: https://my.thechicagoschool.edu/students/Pages/NCADE.aspx

• Email: ncade.me@thechicagoschool.edu

• Phone: (213) 615-7290