When to Use Which Statistical Test

Rachel Lovell, Ph.D., Senior Research Associate
Begun Center for Violence Prevention Research and Education
Jack, Joseph, and Morton Mandel School of Applied Social Sciences

Importance of Using the Appropriate Test

- “Lies, Damned Lies, and Statistics”
- Examples when this goes wrong
- Task made difficult by breadth of statistics

Which Type of Test Is Based on:

- Level of measurement of dependent variable (DV) – including number of DVs
- Level of measurement of the independent variable (IV) - including number of IVs
- Research question(s)
- Statistical assumptions
  - Random sample
  - Normal distribution of the sampling distribution

Levels of Measurement

- Nominal, Ordinal, Interval
- Often organized into categorical (discrete) or continuous (interval)
- “Fuzzy” levels of measurement
  - Ordinal
    - Ex: using Likert scales as interval variable (scales vs. items)
  - Nominal
    - Yes/No or Presence/Absence – must be 0,1
    - Can treat as interval

Research Question(s): Degree of Relationship

- **Correlation**: how much variables co-vary (bivariate and partial)
- **Regression**: prediction of DV based on IV(s) (bivariate and multivariate)
- **Path Analysis** (AKA structural equation modeling): extension of regression – "causal" – what are the direct and indirect effects

Research Question(s):Significance of Group Differences

- **T-test**: group mean differences (interval DV) – 2 groups (IV)
  - *Gender impact SAT scores? (Males and females have different SAT scores?)*
- One-way analysis of variable (**ANOVA**): group mean differences (interval DV) – 2+ groups (IV)
  - Race/ethnicity (2+ grps) impact *SAT scores? (Do their scores significantly differ?)*
- One-way analysis of covariance (**ANCOVA**): like ANOVA except "controlling" for third variable (covariate)
  - *Race/ethnicity impact SAT scores controlling for family income?*
- One-way multivariate analysis of variance (**MANOVA**): >1 DV, controlling for correlations among DVs (assuming correlation among DVs)
  - Race/ethnicity impact reading, math, and overall achievement (DVs) among 6th grade students?

Research Question(s): Significance of Group Differences (con'd)

- One-Way **MANCOVA**: >1 DV while controlling for covariate
  - *Race/ethnicity (IV) impact reading, math, and overall achievement (DVs) among 6th grade students after adjusting for family income?*
- **Factorial MANOVA** – like MANOVA but with 2+ IVs
  - *Race/ethnicity and learning preference* (IVs) *impact reading, math, and overall achievement (DVs) among 6th grade students?*
- **Factorial MANCOVA** – like factorial MANOVA but with >1 covariate
  - *Race/ethnicity and learning preference (IVs) impact reading, math, and overall achievement (DVs) among 6th grade students after controlling for family income?*

Research Question(s): Prediction of Group Membership

- Goal: identify specific IVs that best predict group membership as defined by categorical DV
- **Discriminant Analysis**: combination of interval IVs that best distinguish categorical DV
  - *Which risk-taking behaviors (amt of alcohol use, drug use, sexual activity, and violence - IVs) distinguish suicide attempters from nonattempters - DV?*
- **Logistic Regression:** categorical or interval IVs that predict odds of Y (categorical DV)
  - *Which risk-taking behaviors (amt of alcohol use, drug use, sexual activity, and presence of violence, y/n) increase the odds of a suicide attempt occurring (DV)?*
  - DV can be binary or multinomial

Research Questions: Structure

- Goal: reduce the number of IVs (no DVs)
- **Factor analysis**: explores the underlying structures of scale or instrument (aka latent factors)
  - Exploratory Factor Analysis (EFA) – exploratory stage – not testing theory, let data tell you the factor structure
  - Confirmatory Factor Analysis (CFA) – confirmatory stage – testing theory, theory guides structure, often done after EFA
- **Principal Components Analysis** (PCA): similar to factor analysis, same goal (variable reduction)
  - Usually first exploratory procedure conducted
  - In practice very similar to EFA
  - Preferred method of factor extraction

Level of Measurement: Bivariate

- Categorical IV and DV: **Chi-square**
- Interval DV and categorical IV (2 grps): **independent sample t-test**
- Interval DV and categorical IV (2+ grps): **ANOVA**
- Categorical DV and Interval IV: **discriminant** (if applicable) or can categorize IV and do Chi-square or **logistic** regression
- Interval DV and IV: **correlation** (covariance) and **regression** (prediction)

Level of Measurement: Multivariate

- Categorical IV and DV + categorical covariate: **Chi-square**
- Interval DV and categorical IV + covariate: **ANCOVA**
- Interval DVs and categorical IVs: **MANOVA**
- Other –**OVAS** previously discussed
- Interval IVs and DV: **partial correlation**

Level of Measurement: Multivariate Regression

- Interval IVs and DV
- DV:
  - DV = interval: **multivariate** regression (AKA OLS regression)
  - DV = binary (0,1): **logistic** regression
  - DV = ordinal (ex: 0, 1, 2, 3): **multinomial** regression
  - DV = count (ex: # of children, # visits to doctor in last 6 mths): **negative binomial or Poisson** regression
- IVs: supposed to be interval, if not….
  - Dichotomous (dummy) variable (0,1)
  - Polychotomous (2+ groups): white, black, other – put two in regression, leave one out, coefficients are in reference to the group left out

Statistical Assumptions

- Random sample and normal sampling distribution
- Transformation of nonnormal DV: log, square root, z score
- Nonparametric tests: doesn't assume normal distribution
  - Can use when small sample
  - DV is extremely nonnormal (test for this)
  - More conservative tests; not as statistical "strong" as parametric tests

Common Nonparametric Tests

- Chi-square = nonparametic test
- Pearson's correlation: Spearman's rho and Kendall's tau
- Paired sample t-test: Wilcoxon t test
- Independent sample t-test: Mann Whitney U test
- ANOVA: Kruskall-Wallis test

There are many others ….

Other considerations

- Are your data longitudinal? (more than one observation for same person/observation)
  - Paired sample t-tests (pre and post)
  - Repeated measures ANOVA
  - Regression: fixed, random, and mixed models
- Are your data nested?
  - Multi-level modeling (AKA hierarchical linear modeling)
- Do your data have tons of missing data?
  - Multiple imputations
- Are you interested in survival – how long one stays in a certain state before an event occurs (timing is important)?
  - Survival analysis

Longitudinal Data Analysis (LDA)

- Use when modeling pattern(s) of change in DV over time and need to explore the effects of time-varying predictors and events on individual/unit outcomes
- A more robust model because have more “information” (more variability, less collinearity, more degrees of freedom, and more efficiency)
- 2 points=great; 3 points=awesome; 2 points: change score regression (Allison 1990)

Most Common Models in LDA

- Random Effects (RE) – AKA Error Components Models
  - Ability to control for individual heterogeneity (or variability)
  - Example: Units are same industries observed at multiple time points, DV = racial earning inequality. *Random effect of industry error* assumes that there is some process that creates random variation across industries in the generation of inequality.
- Fixed Effects (FE) – AKA Least Squares with Dummy Variables
  - Ability to control for unobservables (omitted variable bias)
  - Example: Budig, Michelle and Paula England. 2001. "The Wage Penalty for Motherhood." *American Sociological Review*
    - Studying change in women's wages as a result of motherhood. Didn't matter what the unobserved variables impacts wages were because these are held stable (constant). Only interested in what happens to changes over time *within* the individual as a result of motherhood.
- Hybrid of FE and RE – AKA Mixed modeling

Pros and Cons

- Pros: Listed above
- Cons:
  - Complicated modeling (lagged variables, time-varying and invariant variables)
  - Attrition and missing data
  - Must account for and correctly model correlated errors (violation of independence)
  - SPSS doesn't do
  - Getting data in correct format! -- (one of the) most difficult part of LDA: getting data in correct format

Nesting: Hierarchical Linear Model (HLM)

- AKA: multilevel modeling, mixed linear models, random coefficient models
- Used when there is a natural nesting to the data
- Provides coefficients for each levels (parcel out the variance explained by each level)
- Example - DV: student achievement scores; Level 1: student, Level 2: classroom; Level 3: school
  - Must have variables at all levels or can cluster
- Violate assumption of independence – correlated errors, which you MUST correct!

Pros and Cons

Pro:

- Interpretation of the levels very useful

Con:

- Language and formulas are unnecessarily difficult
- Sample size drastically reduced (need any more cases and clusters)
- Can get complex very quickly, especially with the addition of predictors
- SPSS doesn't do

Missing Data

- Mostly concerned when it is on the DV
- Use when:
  - Losing lots of cases
  - Assume that missing cases bias your results
- Three main ways to deal with missing data
  - Ignore: listwise deletion
  - Using ML methods
  - Multiple imputations (can be done in SPSS)

Survival Analysis

- Came from epidemiology, hence, the term survival -- AKA event history analysis
- Use when:
  - Timing is important and must be modeled
  - You have an event (death, job, election, etc.)
- Censoring of data can be a problem
- Difficult:
  - Wrapping your brain about the terminology and logic
  - Determining the correct distribution model (usually Cox models – SPSS can do Cox models)

Structural Equation Modeling (SEM)

- Very good for certain questions: educ, pol sci, psy
- Use when theory of paths is important – direct and indirect paths
- Difficult when: have more than a few (control) variables, data is nonnormal, exploring with lots of variables, have small samples

Steps for Determining Which Test to Use

- Step 1: Identify variables in RQ
- Step 2: Identify variables as IV(s), DV(s), and covariates (aka control variables)
- Step 3: Determine the level of measurement (categorical or interval) for all variables
- Step 4: Determine purpose of the research question – degree of relationship, group differences, etc.
- Step 5: Put all together

Helpful sources of info

- Handout – borrowed heavily from Mertler and Vannatta
- http://www.socialresearchmethods.net/selstat/ssstart.htm
- http://www.microsiris.com/Statistical%20Decision%20Tree/
- https://play.google.com/apps?hl=en

Recommend my “favorite” stats books

Questions?

Contact info:

Rachel Lovell

rachel.lovell@case.edu

216-368-3349

Social Work Leader for 98 Years

Countdown to Centennial

JACK, JOSEPH AND MORTON MANDEL SCHOOL OF APPLIED SOCIAL SCIENCES

CASE WESTERN RESERVE UNIVERSITY