Introduction to Statistics Used in Nursing Research

Laura P. Kimble, PhD, RN, FNP-C, FAAN
Professor and Piedmont Healthcare Endowed Chair in Nursing
Georgia Baptist College of Nursing
Of
Mercer University

Copyright 2013 Laura Kimble. Do not reproduce without permission
Classification of Types of Statistics

• Descriptive Statistics
  – Nominal/ordinal level: data frequencies, cross-tabulation tables
  – Interval/ratio level: frequencies with histograms, measures of central tendency

• Inferential Statistics
  – Interval estimates-confidence intervals
  – Hypothesis Testing
    • Parametric Tests
    • Non-parametric Tests

Copyright 2013 Laura Kimble. Do not reproduce without permission
Parametric Vs. Non-Parametric Tests

- **Parametric Tests**
  - Most common statistics used in research
  - To provide accurate results, data must meet statistical assumptions
  - More powerful, meaning more likely to obtain a statistically significant result, if one exists
  - Ex: t-test, analysis of variance, Pearson correlation

- **Non-parametric Tests**
  - Less commonly used in research
  - Good for use with small sample sizes or data not normally distributed
  - Most parametric tests have a non-parametric equivalent
  - Ex: Mann-Whitney U, Kruskal Wallis, Spearman’s rho correlation
So what is a p value really....

- A probability or how many times out of 100 that what we observed in the study would occur simply by chance in the population

- Words that mean the same thing as p value include: alpha, significance level, risk of Type I error

- The convention in research is to set a significance level prior to the research beginning, usually p < .05

- Then any analysis that has an associated p value of < .05 is considered statistically significant
For example

- A study is conducted to address the following research hypothesis:
  - 1) Children ages 6 to 8 years old with asthma will report significantly poorer physical function than children ages 6 to 8 years old with sickle cell disease.
    - The researcher conducts a t-test and finds that children with asthma report significantly poorer function \((p = .02)\) than children with sickle cell
    - The \(p = .02\) means that 2 times out of 100, the difference observed between the two groups in the population could have occurred by chance, which means 98 times out of 100 it was likely to be a true difference. Since \(p = .02\) is less than \(p = .05\) then the research hypothesis is supported.
Choosing The Appropriate Statistical Test

• This is a KEY area for analysis
• Issue that is challenging for researchers and clinicians
• To figure this out, need to think about the following issues:
  – What is the independent and dependent variable?
  – What is the level of measurement for both the independent and dependent variable?
  – What type of research hypothesis are you answering? For example, differences between independent or related groups, relationships between variables
  – Do the data meet the statistical assumptions of specific test?
Review

• Independent variable-
  – Variable that causes or influences another variable
  – AKA predictor variable
  – In intervention studies, the intervention (experimental vs. control) is always the independent variable when testing the effect of the intervention

• Dependent variable-
  – Variable that is influenced by another variable.
  – Outcome variable
  – AKA criterion variable
Levels of Measurement

- **Nominal** (Categorical): Lowest on scale data organized into categories only
- **Ordinal**: Categories which can be ordered by rank, but intervals between numbers are not truly equal
- **Interval**: Distances between scale are equal, however 0 does not mean absence of the concept
- **Ratio**: Highest form of measurement, equal distance between points on scale and a zero means absence of concept
Research Questions and Statistics Strongly Linked

• How research question is stated or proposed guides what data are to be collected
• The types of variables in the study and how the data are collected impact what statistics can be used
• Main situation to be avoided is to not have adequate data to answer the research question.
Considerations for Research Questions

• Is the topic too broad or too narrow?
  – For example: What is the best way to manage patients with diabetes? (TOO BROAD)

• Is the topic appropriate for research?
  – Do people with assertive personalities make the best nurses? (NOT REALLY RESEARCHABLE)
  – (Any question that involves providing less than the standard of care is not appropriate for research)

• Are the data to be collected for potential publication or more for “in-house” program planning and evaluation?
  – If publication is planned, need to obtain IRB approval before implementation (check with own IRB for required consent/approval)

• Can you operationally define the variables? - meaning can you measure the variables in a way that provides appropriate data
  – Very important, in that you have to be able to measure the variable or thing you are studying. If you cannot measure it, you cannot use statistics to examine it.
Types of Research Questions

• Descriptive research question
  – Univariate (or one variable) research question

• Type I research question
  – Research questions that ask how things are related

• Type II research question
  – Research questions that ask how or if things or groups are different

Descriptive Research Questions

• Example: What is the severity of nausea experienced by patients during treatment with adriamycin and cyclophosphamide chemotherapy regimen?
  
  – Note: univariate (severity of nausea)
  
  – Because univariate, not a question answered with inferential statistics. In other words, there is no p value to report.
Other Examples of Descriptive Research Questions

• What is the *prevalence of falls* among elders admitted to the neurological unit from June 1, 2009 through May 31, 2010?
• What is the *frequency of bowel incontinence* among patients with irritable bowel syndrome?
• Among post-CABG patients, what proportion are *extubated within 6 hours post-operatively*?
• What *percentage* of gastric band procedure patients are male?
Statistics to Answer Descriptive Research Questions

• Use descriptive statistics to address these research questions such as percentages, means and standard deviations, and proportions.

• If data are categorical data, for example: gender or marital status would use percentages. For instruments that are scored, would use means and standard deviations.
Type I Research Questions

• Asks how things are related
• Example research questions:
  – Is there a relationship between age and severity of pain following shoulder replacement surgery?
  – Is greater depression following AMI associated with poorer self management of medications following hospital discharge?
  – Is there a relationship between living arrangements and hospital readmission within 30 days following discharge among CHF patients?
Analyzing Type I Research Questions

• Note for these research questions there are now two variables involved
• Use inferential statistics that measure associations or relationships
• For nominal level data would use chi-square
• For ordinal level data or small sample sizes would use Spearman’s rho (non-parametric statistic)
• For interval/ratio level data would use Pearson correlation
Type II Research Questions

- Research questions that ask how things or groups are different from each other

Example research questions:
- Is there less groin site bleeding in patients using manual compression vs. c-clamp in the first 24 hours following PCI?
- Do Hispanic, White, and African American patients utilize mental health services differently following drug/alcohol inpatient treatment?
- Does intent to stay differ for nurses working in a mental health setting compared to those working in a medical-surgical setting?

Independent variable in red; Dependent variable in green
Analyzing Type II Research Questions

- Note for these research questions there are now two variables involved, but there is more clearly an independent and a dependent variable.
- Use inferential statistics that measure group differences.
- In this case the independent variable is a nominal grouping variable (which may have 2 or more groups depending on the variable).
- The dependent variable may be nominal, ordinal or interval/ratio.
- For nominal level data would use chi-square.
- Comparing 2 groups would use t-tests or Mann Whitney U.
- Comparing greater than 2 groups would use ANOVA or Kruskal Wallis.

Copyright 2013 Laura Kimble. Do not reproduce without permission.
Let’s Take Some Examples- Identify Type of Research Question and the IV (s) and DV(s)

• Are there differences in subjective discomfort among patients receiving tap water, soapsuds or PEG-ES (polyethylene-glycol-electrolyte) enema solutions?
• Is there a relationship between partner verbal abuse and depression among female adolescents?
• What is the frequency of hospital acquired intravenous line infections in the surgical intensive care unit?
• Do hospital A and Hospital B differ in frequency of intravenous line infections in the surgical intensive care unit?
• What is the lived experience of being a colon cancer survivor? (NOTE: THIS IS A TRICK QUESTION)
• Are family members satisfied with the revised policy on family visitation on the pediatric neurology unit?
Questions?