STATISTICAL CONSIDERATIONS IN EDUCATIONAL RESEARCH

Dalia Meisha, DDS, MPH, DScD
dmeisha@vcom.vt.edu
Outline

Before
- Sample Size Calculation
- Research Design
  - Objectives → Study Outcome(s)
  - Research Variables
  - Sampling
  - Reliability
- Statistical Analysis Plan

During
- Data Coding/ Data Collection Forms

After
- Statistical Analysis
VCOM Edu. Research

- Evaluation of E-Mail Board Review Questions by Osteopathic Medical Students – The Weakly Bored
- The Effect of Physical Touch on Patient Satisfaction in a Standardized Patient Model
- Methods of Delivering Public Health Messages to Children in Underserved Areas of Honduras
- A Comparison of Medical Student versus Preceptor Evaluation of Clinical Presentation Skills
- The Virtual Standardized Patient: an Effective Modality for Educating Preclinical Medical Student Presentation Skills. A Comparative Item Analysis of Live Standardized Patients versus Virtual Patients
A Comparative Item Analysis of Live Standardized Patients versus Virtual Patients

Objective: to perform comparative analysis of the quality of the case presentation based on an item analysis of live (Standardized Patients) SP versus VP (virtual patients).

Sample: A total of 340 students

Outcome: Evaluation of students’ performance (4 areas of the case presentation: history, physical assessment and laboratory evaluation, differential, and plan)

Predictor: live SP versus virtual patients

Results: Performance on the virtual case presentations and live patients were found to be comparable (percentage of correct answers was 77.3% and 76.8% respectively, p = 0.9).
Outline

**Before**
- Sample Size Calculation
- Research Design
  - Objectives → Study Outcome(s)
  - Research Variables
  - Sampling
  - Reliability
- Statistical Analysis Plan

**During**
- Data Collection Forms

**After**
- Statistical Analysis
Why Sample Size Calculation?

- **Scientifically:**
  - Too small sample size → study will lack the precision
  - Too small sample size → we may mistakenly conclude no difference (insufficient power, Type II error)

- **Ethically:**
  - Too small sample size → exposed subjects to testing without the capability to advance knowledge
  - Too big sample size → potential to expose unnecessarily large number of subjects (potential harm)

- **Economically:**
  - Too small sample size → waste resources
  - Too big sample size → waste resources
Sample Size Calculation

• Desired level of significance (α)
• Desired power
• Statistics that will be used for analysis
• Whether the test would be one or two-tailed
• Allocation ratio
• Size of difference of main outcome and SD: Literature or Pilot study
• Drop out?
Sample Size Calculation

• Desired level of significance ($\alpha$): 0.05
• Desired power: 80%
• Statistics that will be used for analysis: T-test comparing 2 means
• 2-tailed
• Allocation ratio: 1:1
• Size of difference (94.4, 89.7), SD (6.9, 14.6)
**Power Analysis in G*Power 3.1.7**

**Central and noncentral distributions**

**Protocol of power analyses**

- Critical t = 1.9728

**Test family**
- t tests

**Statistical test**
- Means: Difference between two independent means (two groups)

**Type of power analysis**
- A priori: Compute required sample size - given α, power, and effect size

**Input Parameters**
- Tail(s): Two
- Determine =>
  - Effect size d: 0.4116082
  - α err prob: 0.05
  - Power (1-β err prob): 0.8
  - Allocation ratio N2/N1: 1

**Output Parameters**
- Noncentrality parameter δ: 2.8218436
- Critical t: 1.9728001
- Df: 186
- Sample size group 1: 94
- Sample size group 2: 94
- Total sample size: 188
- Actual power: 0.8015817

**Additional parameters**
- n1 = n2
- Mean group 1: 94.4
- Mean group 2: 89.7
- SD σ within each group: 6.9
- SD σ group 1: 6.9
- SD σ group 2: 14.6

**Buttons**
- Calculate and transfer to main window
- X-Y plot for a range of values
"A sample size of 94 in each group will be sufficient to detect a difference of 4.7 points on students performance, assuming standard deviation of 6.9 and 14.6 points, a power of 80%, and a significance level of 5%.

Option: This number has been increased to 104 per group (total of 208), to account for a predicted drop-out from vcom of around of 10%"
Factors affecting Sample Size Calculation:

• \( \uparrow \) power \( \Rightarrow \) sample size \( \uparrow \)

• \( \downarrow \) variability \( \Rightarrow \) sample size \( \downarrow \)

• \( \uparrow \) Size of differences \( \Rightarrow \) sample size \( \downarrow \)
Pitfalls in Sample Size Calculation

• A sample size of 100 is proposed parallel to a published similar study that recruited 100 subjects & found highly significant results.
• Sample size calculation is not feasible because there is no available information.
  • Do pilot study
• Number was decided based on available subjects
  • Multi-center study
Outline

Before
- Sample Size Calculation
- Research Design
  - RQs/ Objectives → Study Outcome(s)
  - Research Variables
  - Sampling
  - Reliability
- Statistical Analysis Plan

During
- Data Collection Forms

After
- Statistical Analysis
Research Design

Research Question: Is the outcome of the students’ case presentation different with live SP versus virtual patients?

Hypothesis: Students Performance with live SP=with VP

Comparison/ control: Live SP vs Virtual Patients

Outcome (Dependent): Students’” performance

Predictor (Independent ): Live SP vs VP
## Research Variables

<table>
<thead>
<tr>
<th>Outcome (Dependent)</th>
<th>Independent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Influenced by manipulated variable</td>
<td>• Intervention</td>
</tr>
<tr>
<td>• Define the outcome measure</td>
<td>• Demographics</td>
</tr>
<tr>
<td>• Measureable</td>
<td>• Other possible confounders</td>
</tr>
</tbody>
</table>

- Outcome (Dependent): Influenced by manipulated variable, Define the outcome measure, Measureable
- Independent Variables: Intervention, Demographics, Other possible confounders
Research Variables

Continuous

- Students grade
- Age in years

Categorical

- Students letter grade (A, B, C, F)
- Age in categories
Sampling

- Simple Random (randomization table)
- Systematic
- Stratified
Reliability

- Calibration
- Inter-Rater Reliability: with other faculty
- Intra-Rater Reliability: same faculty at another occasion

http://www.socialresearchmethods.net/kb/reltypes.php
Outline

Before
- Sample Size Calculation
- Research Design
  - RQs/Objectives \rightarrow Study Outcome(s)
  - Research Variables
  - Sampling
  - Reliability
- **Statistical Analysis Plan**

During
- Data Collection Forms

After
- Statistical Analysis
To convince the reviewers that you can analyze the data

- Shows awareness of statistical aspects and capability to address issues is important
Outline

Before
- Sample Size Calculation
- Research Design
- RQs/ Objectives → Study Outcome(s)
- Research Variables
- Sampling
- Reliability
- Statistical Analysis Plan

During
- Data Coding/ Data Collection Forms

After
- Statistical Analysis
**Data Coding**

- **What is Data Coding?**
  “The process by which verbal data are converted into variables and categories of variables using numbers, so that the data can be entered into computers for analysis.”

**Variables:**
- Gender
  - Male = 1
  - Female = 2

**VCOM Class**
- 2016 = 1
- 2017 = 2
- 2018 = 3

---

Bourque, Linda B. "Coding." In The Sage Encyclopedia of Social Science Research Methods, 2004
## Data Collection Form

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>AA</th>
<th>AB</th>
<th>AC</th>
<th>AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Trial Number</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
</tr>
<tr>
<td></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
</tr>
<tr>
<td></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
<tr>
<td>2</td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
</tr>
<tr>
<td>3</td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
<tr>
<td>4</td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
</tr>
<tr>
<td>5</td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
<tr>
<td>6</td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
</tr>
<tr>
<td>7</td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
<tr>
<td>8</td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
</tr>
<tr>
<td>9</td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
<tr>
<td>10</td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
</tr>
<tr>
<td>11</td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
<tr>
<td>12</td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
</tr>
<tr>
<td>13</td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
<td><strong>(2) Somewhat confident</strong></td>
<td><strong>(3) Neutral</strong></td>
<td><strong>(4) Somewhat unconfident</strong></td>
<td><strong>(5) Very unconfident</strong></td>
<td><strong>(1) Very confident</strong></td>
</tr>
</tbody>
</table>

- **Do you believe training in mobile dental clinics improved your cultural competency skills as general dentist?**

  - **(1) Yes**
  - **(2) No**
  - **(3) Somewhat**
Data Collection Form
Outline

Before
- Sample Size Calculation
- Research Design
  - RQs/ Objectives → Study Outcome(s)
  - Research Variables
  - Sampling
  - Reliability
  - Statistical Analysis Plan

During
- Data Coding/ Data Collection Forms

After
- Statistical Analysis
Statistical significance (p-value)

• Relates to the probability of rejecting the null hypothesis and accepting the alternative hypothesis

• Significance is shown at \( \alpha 0.05 \)
  • \( p \leq 0.05 \): results occurring are due to chance are 5% or less
Type I and II errors

Type I error ($\alpha$):
- Rejected the null hypothesis when it is true
- Concluded there is a difference between the means of the two groups when, in fact, there is not a difference

Type II error ($\beta$):
- Accepted the null hypothesis when it is false
- Concluded there is no difference between the means of the two groups when in fact there is a real difference

Type I and II errors

<table>
<thead>
<tr>
<th>Conclusion</th>
<th>No effect</th>
<th>Has an effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>Type I error</td>
<td>✓</td>
</tr>
<tr>
<td>Fail to reject $H_0$</td>
<td>✓</td>
<td>Type II error</td>
</tr>
</tbody>
</table>
How to determine statistical test?

- What are you interested in?
- Is your dependent variable Continuous or Categorical
- Data: Parametric versus non-Parametric
### How to determine statistical test?

<table>
<thead>
<tr>
<th>Goal</th>
<th>Type of Data</th>
<th>Statistical Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe one group</td>
<td>Measurement</td>
<td><strong>t</strong> test or <strong>Wilcoxon test</strong> or <strong>Chi-square test</strong></td>
</tr>
<tr>
<td>Compare one group to a hypothetical value</td>
<td>Rank, Score, or Measurement</td>
<td><strong>Chi-square test</strong> or <strong>Binomial test</strong></td>
</tr>
<tr>
<td>Compare two unpaired groups</td>
<td>Binomial (two possible outcomes)</td>
<td><strong>Log-rank test</strong> or <strong>Mantel-Haenszel</strong></td>
</tr>
<tr>
<td>Compare two paired groups</td>
<td>Survival Time</td>
<td><strong>Kaplan-Meier survival curve</strong></td>
</tr>
<tr>
<td>Compare three or more unmatch groups</td>
<td><strong>t</strong> test</td>
<td><strong>MCNEMAR'S TEST</strong></td>
</tr>
<tr>
<td>Compare three or more matched groups</td>
<td>Repeated-measures ANOVA</td>
<td><strong>Friedman test</strong></td>
</tr>
<tr>
<td>Quantify association between two variables</td>
<td>Pearson correlation</td>
<td><strong>Cochran's Q</strong></td>
</tr>
<tr>
<td>Predict value from another measured variable</td>
<td>Simple linear regression or Nonlinear regression</td>
<td><strong>Simple logistic regression</strong></td>
</tr>
<tr>
<td>Predict value from several measured or binomial variables</td>
<td><strong>Multiple linear regression</strong> or <strong>Multiple nonlinear regression</strong></td>
<td><strong>Cox proportional hazard regression</strong></td>
</tr>
</tbody>
</table>
Common Mistakes

- Unclear Subject Recruitment
- Sample Size
- Information about Group Comparison
- Appropriate Statistical Methodology (example: Comparing pre- and post-study Results)
- Reporting Statistical Significance
- Justification for Research
- Distribution of a Variable
- Unclear Research Methods
- Trimming and Cooking Data
- Using Power Analysis to Determine the Appropriate Number of Subjects
THANK YOU