NOVA SOUTHEASTERN UNIVERSITY
Abraham S. Fischler School of Education

Module 4: Inferential Statistics

The Applied Research Center

Module 4 Overview

- Inferential Statistics
- Brief Introduction to Probabilities
- Hypothesis Testing

Parameter vs. Statistic

- A population is the **entire** set of individuals that we are interested in studying
- A sample is simply a subset of individuals selected from the population
- In most studies, we wish to quantify some characteristic of the population → parameter
- Parameters are generally unknown, and must be estimated from a sample
- The sample estimate is called a statistic

Inferential Statistics

- Techniques that allow us to make inferences about a population based on data that we gather from a sample
- Study results will vary from sample to sample strictly due to random chance (i.e., sampling error)
- Inferential statistics allow us to determine how likely it is to obtain a set of results from a single sample
- This is also known as testing for "statistical significance"

Statistical Significance

- Consider a small weight loss study of 40 patients.

- After such a study is over, we want to make generalizations about a larger group (e.g. all similar patients in the nation), but, since it is a small study, the results will be inexact.

- Statistical significance helps us by giving us a "ballpark range" (i.e., confidence interval) around the number (for example the amount of weight lost), encompassing the true number.

Statistical Significance (cont'd)

- If the range is small enough ($p < .05$), we say we are confident that the true amount of weight lost is "more than zero" and "statistically significant."

- Naturally, it says nothing about the practical significance, since the patients might have lost just a gram of weight!

Statistical Significance Testing

Hypothesis Testing

A Brief Introduction to Probability

- A basic understanding of probability is needed
- The probability of an outcome (A), can be thought of as a fraction, proportion, or percentage

$$\text{probability of A} = \frac{\text{number of A outcomes}}{\text{total number of outcomes}}$$

Probability Examples

- What is the probability of rolling a single die and coming up with a six?
  - There is only 1 outcome A (a six)
  - There are 6 possible outcomes (1 to 6)
  - The probability is 1/6 = .1667 = 16.7%
- What is the probability of obtaining a red number in the game of roulette?
  - There are 18 red numbers (A)
  - There are 38 numbers total
  - The probability is 18/38 = .4737 = 47.4%

The Null Hypothesis

- The null hypothesis **always states** that nothing is going on
  - There is no difference, no relationship, no treatment effect, etc.
  - H_0: X = Y
- The alternate hypothesis states that there is a difference
  - H_a: X ≠ Y (non directional)
  - H_a: X > Y or Ha: X < Y (directional)

The (Somewhat Twisted) Logic of Significance Testing

- Compute a probability value that tells how likely our data (or results) would occur just by chance

- If the probability is "low" (e.g., $p = .02$), this means our data is inconsistent with the null
  - There is evidence that there is a difference
- If the probability value is "high" (e.g., $p = .30$), this means our data is consistent with the null
  - There does not seem to be evidence that there is a difference

More on the Logic

- The confusing thing is that we are not directly testing whether or not there is a treatment effect, or relationship
- We are testing how consistent the data is with the hypothesis that there is **no** treatment effect, relationship, etc.
- Thus, a treatment effect is demonstrated indirectly if the data is inconsistent with the null hypothesis

Rule of Thumb ($p < .05$)

- How inconsistent with the null does the data have to be to demonstrate an effect?
- Conventional rules use a $p < .05$ cutoff
- If the data yields a probability value less than .05 ($p < .05$), that means the data is inconsistent with the null, which states no treatment effect or relationship exists ($H_0: X = Y$)
- Therefore, we reject the null

Two Outcomes

- If $p < .05$, our data is inconsistent with the null
  - We "reject the null" and declare our results "statistically significant"
- If $p > .05$, our data is consistent with the null
  - We "fail to reject the null" and declare our results "statistically non-significant"

Example 1

- Suppose we were comparing how males and females differed with respect to their satisfaction with an online course
- The null hypothesis states that men (X) and women (Y) do not differ in their levels of satisfaction
  - $H_0: X = Y$

Example 1 (cont' d)

- On a 25-point satisfaction scale, men and women differed by about 5 points (means were 18.75 and 23.5, respectively)
- They were not identical, but how likely is a 5 point difference to occur just by chance?

Example 1 (cont'd)

- An analysis was conducted, and the p-value for the gender comparison was p = .11
- Thus, there was about a 11% chance that this data (the 5 point difference) would occur by chance
- The p-value is greater than .05, so we would fail to reject the null (results are not significant)
- Thus, there is no evidence that males and females differ in their satisfaction

Example 2

- Suppose we were comparing how males and females differed with respect to how likely they would be to recommend an online course (measured on a 5 point scale)
- The null hypothesis states that there is no difference between men and women in their recommendation of an online course.
  - (H_0: X = Y)

Example 2 (cont’d)

- On a 5-point satisfaction scale, men and women differed by about 1 point (means were 4.3 and 3.1, respectively)
- They were not identical, but how likely is a 1 point difference to occur by chance?

Example 2 (cont' d)

- An analysis was conducted, and the p-value for the gender comparison was $p = .03$
- Thus, there was only a 3% probability that this data would occur by chance
- The p-value is less than .05, so we would reject the null (results are significant)
- Thus, there is evidence that males and females differ in their recommendations

The Meaning of Statistical Significance

- p-values tell how likely it was that our sample was drawn from a hypothetical population where "nothing was going on"
- Thus, the term "statistical significance" simply means that the obtained results are unlikely to represent a situation where there was no relationship between variables
- The difference is big enough to be unlikely to have happened simply due to chance

Cautionary Note

- Just because results are statistically significant, does not mean that the results are of practical importance
- It ends up that large samples are more likely to yield "significant results", even if the differences are rather trivial
- Don't equate "statistical significance" with a "large" or "important" effect

Module 4 Summary

- Inferential Statistics
- Brief Introduction to Probabilities
- Hypothesis Testing

Review Activity and Quiz

- Please complete the Module 4 Review Activity: Hypothesis Testing located in Module 4.
- Upon completion of the Review Activity, please complete the Module 4 Quiz.
- Please note that all modules in this course build on one another; as a result, completion of the Module 4 Review Activity and Module 4 Quiz are required before moving on to Module 4.
- You can complete the review activities and quizzes as many times as you like.

Upcoming Modules

- Module 1: Introduction to Statistics
- Module 2: Introduction to SPSS
- Module 3: Descriptive Statistics
- Module 4: Inferential Statistics
- Module 5: Correlation
- Module 6: *t*-*T*ests
- Module 7: ANOVAs
- Module 8: Linear Regression
- Module 9: Nonparametric Procedures