CH. 7 – Introduction to Hypothesis Testing

PY 221 Research Methods & Statistics I

Dr. Valenti

Outline for Ch. 7

1. Underlying logic of null hypothesis significance testing (NHST)
2. Steps of NHST
3. Common misunderstandings and criticisms related to NHST
4. Effect sizes
5. Using confidence intervals for hypothesis testing (Ch. 8 pp. 154-158)
6. Types of Errors when engaging in NHST

First, some definitions . . .

- A *hypothesis* is a specific prediction about the variables involved in a phenomenon that researchers are studying

- A *hypothesis test* is a statistical method that uses sample data to evaluate a hypothesis about a population

Null hypothesis significance testing (NHST)
underlying logic

- **How probable is it that our sample**, which has particular characteristics (i.e., sample statistics, x-bar), **came from a specific, known population**?

- *Remember:* **Even a *random* sample drawn from a population will always look a little different, have slightly different characteristics, than the actual population**

Null hypothesis significance testing (NHST)
underlying logic

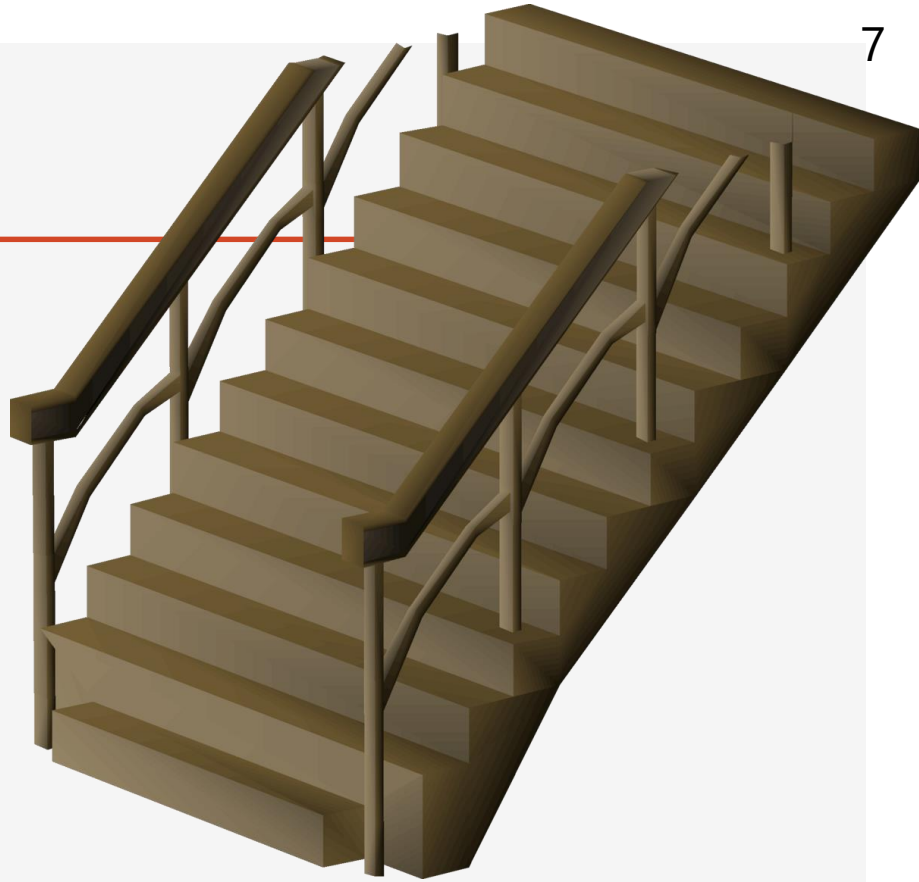
- My sample could look different from **population *A*** . . .
  1. due to chance (*although my sample really does come from* ***population A***)
  2. because my sample actually came from **population *B***

- NHST gives us criteria so that we can decide which of these options is more likely to be true.

Outline for Ch. 7

1. Underlying logic of null hypothesis significance testing (NHST)
2. **Steps of NHST**
3. Common misunderstandings and criticisms related to NHST
4. Effect sizes
5. Using confidence intervals for hypothesis testing (Ch. 8 pp. 154-158)
6. Types of Errors when engaging in NHST

BRIEF OVERVIEW OF STEPS of NHST



1. State 2 hypotheses about unknown population
2. Set your significance level
3. Calculate the test statistic & p-value
4. Make the decision

Example study on cognitive functioning and antioxidants

- Background literature: Aging causes a decline in cognitive functioning (Bartus, 1990)
- Research question: *Does eating blueberry supplements (which have antioxidants) change this decline?*
  - Study Method:
    - Choose cognitive test for which older adults typically score $\mu = 80$ ($\sigma=20$).
    - Randomly select 40 older adults, give blueberry supplements for 6 mo.
    - Give the cognitive test to these 40 adults after 6 mo. of supplements

If we collected these data, how would we know whether the blueberry supplements had an effect on test scores?

→ *Use null hypothesis significance testing*

NHST

Step 1: State 2 hypotheses about the unknown population

- **Null hypothesis, H_0**
  - *Defined conceptually*: In the unknown population, there is no change, no difference, or no relationship; IV has no effect on DV.
- **Alternative hypothesis, H_1 (your book uses H_A)**
  - *Defined conceptually*: In the unknown population, there *is* a change, a difference, or a relationship; IV *does* have an effect on DV.

TO KEEP IN MIND: Since researchers typically predict that their variables **WILL** show a change/difference/relationship/effect, researchers are hoping their sample data **DISCREDITS** (or **casts doubt on**) the **NULL** hypothesis.

NHST

Step 1: State 2 hypotheses about the unknown population

- **Null hypothesis, H_0**
  - *Defined conceptually*: In the unknown population, there is no change, no difference, or no relationship; IV has no effect on DV.
- **Alternative hypothesis, H_1**
  - *Defined conceptually*: In the unknown population, there *is* **a** change, **a** difference, or **a** relationship; IV *does* have **an effect** on DV.
- In this class (as in most psychology research), we will deal only with ***non-directional*** **alternative hypotheses.**

NHST Step 1: State 2 hypotheses about the unknown population

- **Null hypothesis, H_0**
  - *Conceptually*: In the unknown population there is no change, no difference, or no relationship; IV has no effect on DV
  - H_0: $\mu_{b\text{-}s}$ = ?

($\mu_{b\text{-}s}$ refers to the mean cognitive test score of the population of blueberry-supplemented older adults)

NHST Step 1: State 2 hypotheses about the unknown population

- **Null hypothesis, H_0**
  - *Conceptually*: In the unknown population there is no change, no difference, or no relationship; IV has no effect on DV
  - H_0: $\mu_{b\text{-}s} = 80$

    in words, The mean test score for blueberry-supplemented (b-s) older adults is 80

NHST Step 1: State 2 hypotheses about the unknown population

- **Null hypothesis, H_0**
  - *Conceptually*: In the unknown population there is no change, no difference, or no relationship; IV has no effect on DV
  - H_0: $\mu_{b\text{-}s} = 80$

    in words, The mean test score for blueberry-supplemented (b-s) older adults is 80
- **Alternative hypothesis, H_1**
  - *Conceptually:* In the unknown population there *is* a change, a difference, or a relationship; IV does have an effect on DV
  - H_1: $\mu_{b\text{-}s} \neq 80$

  In words, The mean test score for B-S older adults is different from 80

Practice stating the null & alternative hypotheses

Johnson, Tuomisto, & Patching (2011) wanted to know if stress that is created and measured in the *lab* was higher than stress in a *natural setting*.

They plan to compare a measure of physiological stress (heart rate) in people who have their stress created and measured in the lab with people whose stress is created and measured in a natural setting.

1. What is the null hypothesis in words?
2. What is the alternative hypothesis in words?
3. What are the null and alternative hypotheses in symbols?

Practice stating the null & alternative hypotheses

Johnson, Tuomisto, & Patching (2011) wanted to know if stress created and measured in the lab was higher than stress in a natural setting. They plan to compare a measure of physiological stress (heart rate) in people who have their stress created and measured in the lab with people whose stress is created and measured in a natural setting.

1. **What is the null hypothesis in words?**

   There will be **no difference** in heart rates when stress is created and measured in the lab and when it is created and measured in a natural setting.

Practice stating the null & alternative hypotheses

Johnson, Tuomisto, & Patching (2011) wanted to know if stress created and measured in the lab was higher than stress in a natural setting. They plan to compare a measure of physiological stress (heart rate) in people who have their stress created and measured in the lab with people whose stress is created and measured in a natural setting.

1. **What is *another way to write* the null hypothesis in words?**

   Heart rates of participants who are stressed in the lab will be **the same as** heart rates of participants stressed in a natural setting.

Practice stating the null & alternative hypotheses

Johnson, Tuomisto, & Patching (2011) wanted to know if stress created and measured in the lab was higher than stress in a natural setting. They plan to compare a measure of physiological stress (heart rate) in people who have their stress created and measured in the lab with people whose stress is created and measured in a natural setting.

2. **What is the *alternative* hypothesis in words?**

   Heart rates of participants stressed in the lab **will be different** than heart rates of participants stressed in a natural setting.

Another option:

   There **will be a difference** in heart rate when stress is created and measured in the lab and when it is created and measured in a natural setting.

Practice stating the null & alternative hypotheses

Johnson, Tuomisto, & Patching (2011) wanted to know if stress created and measured in the lab was higher than stress in a natural setting. They plan to compare a measure of physiological stress (heart rate) in people who have their stress created and measured in the lab with people whose stress is created and measured in a natural setting.

3. **What are the null and alternative hypotheses in symbols?**

Practice stating the null & alternative hypotheses

Johnson, Tuomisto, & Patching (2011) wanted to know if stress created and measured in the lab was higher than stress in a natural setting. They plan to compare a measure of physiological stress (heart rate) in people who have their stress created and measured in the lab with people whose stress is created and measured in a natural setting.

3. **What are the null and alternative hypotheses in symbols?**

H_0: $\mu_{lab} = \mu_{natural\ setting}$

H_1: $\mu_{lab} \neq \mu_{natural\ setting}$

(assume that μ = the mean heart rate)

Complete three practice questions

Q1: T/F – When engaging in NHST, we're asking the question *How probable is it that our sample came from a particular population?*

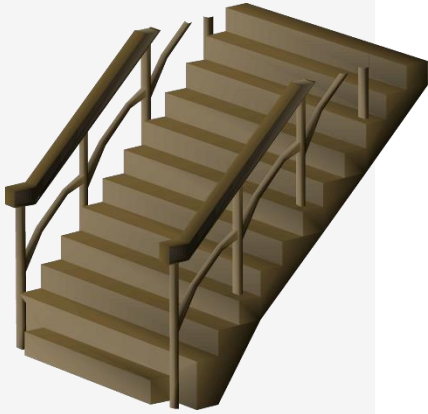
Q2: T/F - In NHST, researchers predict and hope that their sample data casts doubt on the null hypothesis.

Q3: Which is an example of a non-directional alternative hypothesis?

A. Listening to a stats lecture increases your heart rate.
B. Listening to a stats lecture decreases your heart rate.
C. Listening to a stats lecture has no effect on your heart rate.
D. Listening to a stats lecture has an effect on your heart rate.

OVERVIEW of STEPS of NHST

1. State 2 hypotheses about unknown population (null & alternative)
2. **Set your significance level (threshold for rejecting null)**
3. Calculate the test statistic & p-value
4. Make the decision to retain or reject null



- Before collecting data, need to decide what counts as good evidence for discrediting or rejecting the null hypothesis.
- **Think critically: Why should we do this BEFORE collecting data?**
  - *It allows for a fairer, more objective, less biased test of our hypothesis.*

NHST involves asking the question . . .

- **Assuming the null hypothesis were true (credible)**, what's the likelihood that we'd obtain the observed data (i.e., these sample statistics) or data more extreme?

H_0: $\mu_{b\text{-}s} = 80$

- **Assuming the mean test score in the b-s population were 80**, what's the likelihood that we'd obtain a sample with a mean test score of _____ or something more extreme?

Ex: Observed data

$\bar{X} = 79.2$
Ans: High likelihood
(*null seems very credible*)

Ex: Observed data

$\bar{X} = 80.1$
Ans: Very high likelihood (*null seems extremely credible*)

Ex: Observed data

$\bar{X} = 81.5$
Ans: High-ish likelihood
(*null seems very credible*)