Two types of t-tests

review

The two ways of assigning IV conditions for experiments are associated with two different types of t-tests.

- Within-subjects design → Paired samples t-test (repeated measures)
  - *Same* entities are in *all* conditions (i.e., at each level)
    - or entities are naturally *paired* & complete the same measures

- Between-subjects design → Independent samples t-test
  - *Different* entities in each condition (i.e., at each level)
    - or different entities in each naturally-occurring group

Paired samples t-test: What it is used for

- Tests for differences between two groups/conditions when the *same entities* are in each group/condition (or when we have *pairs* of scores for each unit).
  - Ex 1: For the first ½ of semester, students took notes by hand and for the second ½, students took notes using laptops. Each student experienced both note-taking methods. Midterm exam scores were compared with final exam scores.

  - Ex 2: 20 sets (*pairs*) of heterosexual romantic couples were asked how satisfied they were with their relationships, on a scale from 1- not at all *to* 10 – an extreme amount. Scores were compared across males and females to see if they tended to be equally satisfied.

Independent samples t-test: What it is used for

- Tests for differences between two groups when *different (separate, independent) entities* are in each condition (or are in each naturally-occurring group). No pairings are present.
  - Ex 1: Do adults with a college degree differ on blood pressure from adults without a college degree? Measure BP for both and compare.
  - Ex 2: Do PY majors and SO majors differ on graduating GPA? Measure GPA for each group and compare.
  - Ex 3: Invisibility cloak study with a between-subjects design. Ps wear either cloak or no cloak, and mischievous behavior is measured.

Outline for Ch. 9-10 – t-tests

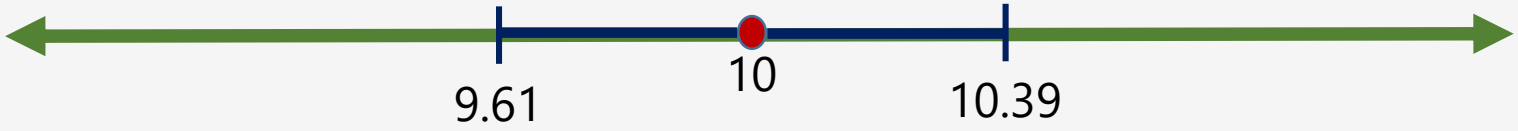
1. Overview of when to use t-tests
2. What is a *t*-test, conceptually?
   - Null and alternative hypotheses
3. Experimental designs & types of t-tests
   - Paired samples *t*-test (aka, repeated measures) - what it's used for
   - Independent samples *t*-test - what it's used for
4. **Confidence intervals for difference scores**
5. Effect sizes

review

Review: What is meant by a CI when estimating a population mean?

- a "95% confidence interval around the mean" translates to . . .
  - *we are 95% confident that the interval (the boundaries) contains the true population mean.*
- point estimate = the sample mean

- 95% CI [9.61, 10.39]



- With a t-test, we're not simply trying to estimate *the mean score of a variable in a population*. Instead, we're trying to estimate . . .
  - ***the size of the difference between two groups' means***

What is meant by a CI when estimating the average size of the difference between the two groups?

- a "95% confidence interval around the mean difference" translates to…
  - *we are 95% confident that the interval contains the true size of the difference in means between the groups in the population.*

- point estimate = the difference score (i.e., the difference in the means of the 2 groups in the sample)

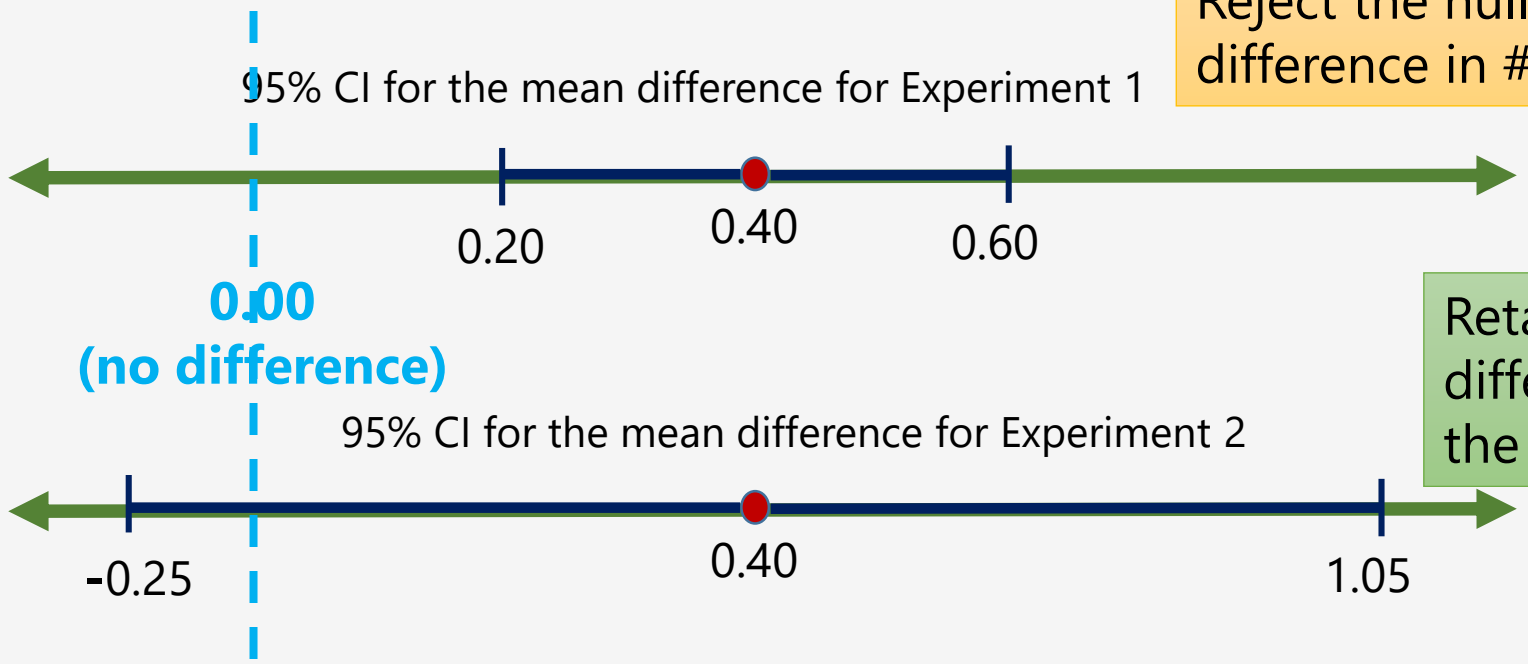
EXAMPLE:
- Group 1 (cloak): **3.50** mischievous acts, on average
- Group 2 (no cloak): **3.10** mischievous acts, on average

What would be the value of our point estimate?
Difference score = **0.40** (or – 0.40, depending on which group you subtract from the other)

How to interpret CIs for mean differences?
How do we know whether to reject or retain the null hypothesis?

- Which experiment's data show evidence for a significant difference btwn the mean # of mischievous acts of those wearing a cloak and the mean # of mischievous acts of those *not* wearing a cloak?

hypothetical sample data



Reject the null and conclude there *is* a significant difference in # of mischievous acts across groups.

Retain the null. Null is credible. The difference in mischievousness between the two groups could in fact be zero.

RECALL: 95% CI for t-test = we are 95% confident that the interval contains the true **size of the difference** in means between the groups in the population.

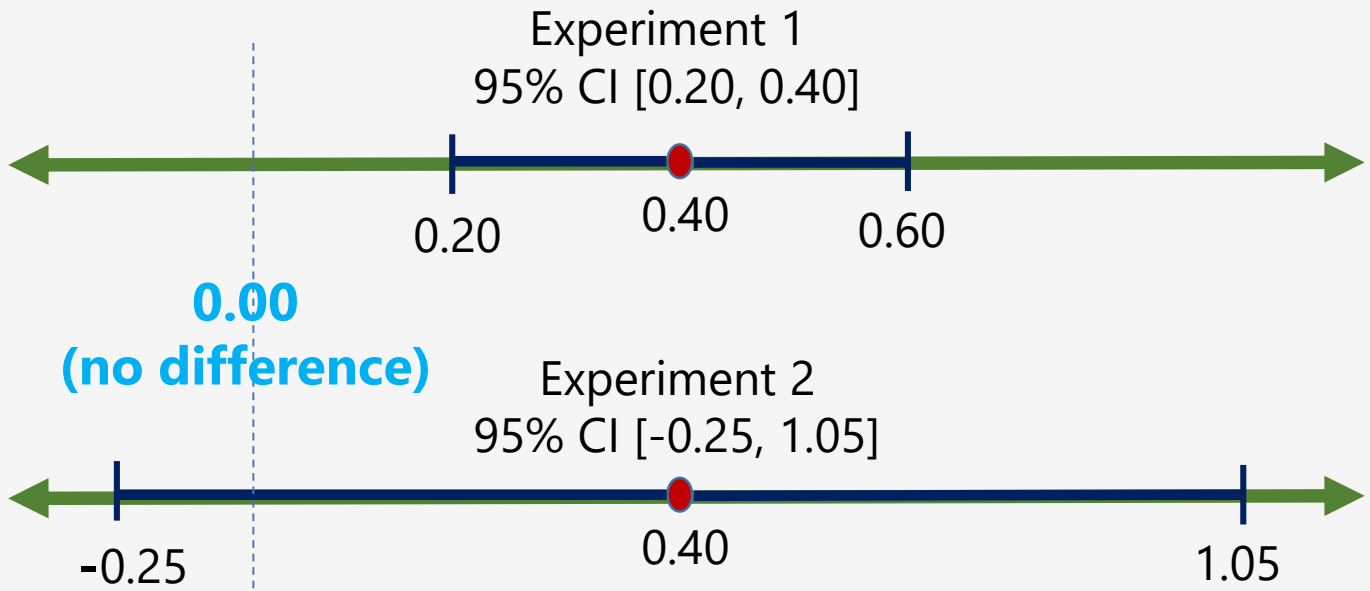
H_0: $\mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

How to interpret CIs for mean differences?

Recall the null & alternative hypotheses
H_0: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$ (difference is zero)
H_1: $\mu_1 \neq \mu_2$ or $\mu_1 - \mu_2 \neq 0$ (difference is not zero)

Rule of thumb for interpretation of CIs for t-tests:

- If CI *contains zero* → *retain* the null hypothesis of no/zero differences (Exp 2)
- If CI doesn't contain zero → reject the null hypothesis and conclude there are significant differences between groups; it's unlikely the difference is 0 (Exp 1)



Practice with confidence intervals around mean differences

1. Which of the following 95% CIs demonstrate a *significant difference* in the mean heights of males and females (in inches)? Select all that apply.

a) 95% CI [0.01, 7.99]
b) 95% CI [-0.06, 0.99]
c) 95% CI [-2.40, -0.25]
d) 95% CI [3.59, 8.05]
e) 95% CI [-2.25, 1.95]

Suppose the mean GPA difference between freshman & seniors in your sample is 0.50 GPA pts. Given this info, do the following:

2. generate a hypothetical CI that would be interpreted as showing NO significant differences between the two groups' GPAs.
3. generate a hypothetical CI that would be interpreted as showing a *significant* difference between the two groups' GPAs.

There are many possible correct answers for Q2 and Q3.

Practice with confidence intervals around mean differences

1. Which of the following 95% CIs demonstrate a *significant difference* in the mean heights of males and females (in inches)? Select all that apply.

a) 95% CI [0.01, 7.99]
b) 95% CI [-0.06, 0.99]
c) 95% CI [-2.40, -0.25]
d) 95% CI [3.59, 8.05]
e) 95% CI [-2.25, 1.95]

Any interval that *doesn't* contain zero suggests a significant difference in heights.

Suppose the mean GPA difference between freshman & seniors in your sample is 0.50. Given this info, do the following:

2. generate a hypothetical CI that would be interpreted as showing NO significant differences between the two groups' GPAs.
3. generate a hypothetical CI that would be interpreted as showing a *significant* difference between the two groups' GPAs.

#2 - Many possible answers. Need to make sure that 0.50 (the mean difference, the pt estimate) is at center of interval, ***and*** make sure interval ***contains 0***.
EX: 95% CI [-0.50, 1.50] or 95% CI [-0.10, 1.10]

#3 - Many possible answers. Need to make sure that 0.50 is at center of interval, ***and*** that interval ***does not*** contain 0. EX: 95% CI [0.45, 0.55] or 95% CI [0.05, 0.95]

Outline for Ch. 9-10 – t-tests

1. Overview of when to use t-tests
2. What is a *t*-test, conceptually?
   - Null and alternative hypotheses
3. Experimental designs & types of t-tests
   - Independent samples *t*-test - what it's used for
   - Paired samples *t*-test (aka, repeated measures) - what it's used for
4. Confidence intervals for difference scores
5. **Effect sizes**

Effect size for t-tests

review

- The *statistical significance* of the difference btwn groups is distinct from the *magnitude* of the difference btwn groups (i.e., the effect size)

- Recall the key added benefit of effect size measures (over p-values):
  - By estimating the *size* of the difference (or change), the ***effect size*** reveals ***practical significance*** of the difference (or change) between the groups

Effect size for t-tests (for both *independent* and *paired samples*)

- Researchers typically only report effect size if the t-test reveals a significant difference between the groups (i.e., $p < .05$)
- Many ways to calculate effect size, but **Cohen's *d*** is best to use for t-tests (JAMOVI can calculate this for us)
- How do we interpret the Cohen's *d* value?
  - Rules of thumb:
    - *d* = +/- .20 ~ "small effect"
    - *d* = +/- .50 ~ "medium effect"
    - *d* = +/- .80 ~ "large effect"

https://rpsychologist.com/d3/cohend/

In this class, we'll use these rules of thumb to evaluate the importance of the effect/relationship/change/etc.

In reality, researchers consider these rules of thumb *alongside* other info…
e.g., the results of other studies on the same topic.
e.g., the relative costs and benefits of applying the results to real life