In case you forgot

Matched Pairs design is a common experimental design for comparing two treatments that uses blocks of size 2

Today

What happens when we have paired data?

Construct Confidence Intervals

Perform a significance test

When to use differences of two means (or p) and when to use procedures for paired data?

10.3 Day 1

10.3

tight for time

There are two ways that a statistical study involving a single quantitative variable can yield paired data:

1. Researchers can record two values of the variable for each individual. (experiment investigating whether music helps or hinders learning)

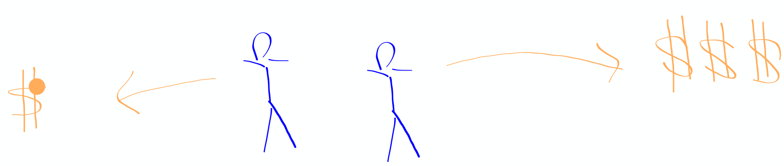
2. The researcher can form pairs of similar individuals and record the value of the variable once for each individual. (observational study of identical twins' IQ scores)

Interesting to Note:

If a random sample of twins is selected, the results can be generalized to all identical twins

but..

Researches could not infer a cause and effect relationship between family income and twins' IQ because random assignment was not passsible



Proper analysis looks at the differences
(not as two separate populations)

**Analyzing Paired Data**

To analyze paired data, start by computing the difference for each pair.

Then make a graph of the differences. Use the mean difference $\bar{x}_{diff}$ and thestandard deviation of the differences $s_{\text{diff}}$ as summary statistics.

therefore conditions
for inference
have to focus
on the difference

i.e. $\mu_{diff}$

10.3 Day 1 **Does Memory Training Help?**


Does memory training help? You will be given a list of words. You will be given 3 minutes to memorize as many as possible by just rereading the words using memorization strategy A. You will record as many words as you can remember. You will then be given another list of words to memorize for 3 minutes using a memorization strategy B. You will record as many words as you can remember.

let's read together
once you have the handout

Look at Strategy A
( don't look at the back, yet )

✓ I'll time you for 3 minutes.

✓ When I say stop, write as many Words as you can on the scratch paper.

✓ Then we'll repeat for Strategy B a different set of words.

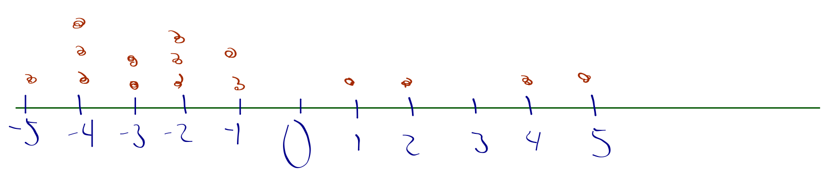
be sure you record the correct order.

3. Add your *paired data* to the table on the board. Copy the data below and calculate the difference in each pair.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| A | | | | | | | | | |
| B | | | | | | | | | |
| Difference | 3 | 1 | 2 | -2 | 2 | -1 | -3 | -4 | 4 |

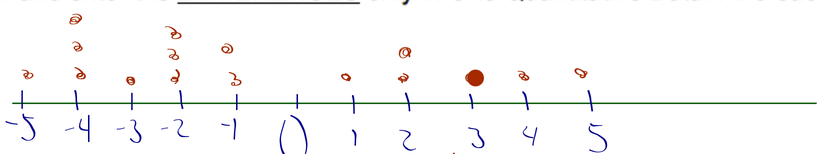
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| A | | | | | | | | | |
| B | | | | | | | | | |
| Difference | -2 | 5 | -2 | -5 | -4 | -4 | -1 | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| A | | | | | | | | | |
| B | | | | | | | | | |
| Difference | | | | | | | | | |



**4.** Go to stapplet.com and enter the <u>DIFFERENCES</u> only in One Quantitative Data. Find each of the following:

Dotplot:




DIFFerence • is #words

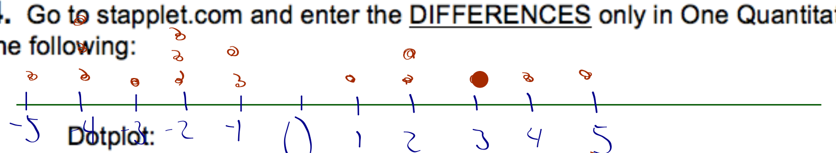
What does the dotplot suggest about the memory training?

Mean: ____________ Interpret:

Standard Deviation: _______

**4.** Go to stapplet.com and enter the <u>DIFFERENCES</u> only in One Quantitative Data. Find each of the following:

Dotplot:




DIFFerence • is #words (A - B)

What does the dotplot suggest about the memory training?

Suggests Strategy B is a bit better

← Identify where most dots are located and tell which strategy is

Mean: -0.69 Interpret:

$\bar{x}_{diff}$ =

Standard Deviation: 3.14

$S_{diff}$

"On average, people remembered .7 words more words with strategy B."

5. Confidence Intervals for $\mu_{diff}$

mean difference

**5.** Construct a 95% confidence interval for the true mean difference in words remembered by students using rereading and story.

**STATE: State the parameter you want to estimate and the confidence level.**

Parameter: $\mu_{diff}$ → the mean difference in words remembered (A − B)

Statistic: $\bar{x}_{diff} = -.69$

Confidence level: 95% CI

**PLAN: Identify the appropriate inference method and check conditions.**

Name of procedure: One-sample t interval for $\mu_{diff}$

Check conditions: Rand

10%

Normal/Large Sample

therefore conditions for inference have to focus on the difference → for Confidence Intervals calculations are just a one-sample t interval for a mean (also called a paired t interval)

Ie. $\mu_{diff}$

**5.** Construct a 95% confidence interval for the true mean difference in words remembered by students using rereading and story.

**STATE: State the parameter you want to estimate and the confidence level.**

Parameter: $\mu_{diff}$ → the mean difference in words remembered

Statistic: $\bar{x}_{diff} =$

Confidence level: 95% CI

**PLAN: Identify the appropriate inference method and check conditions.**

Name of procedure: One-sample t interval for $\mu_{diff}$

Check conditions: Rand "randomly assigned"

~~10%~~ N/A

Normal/Large Sample
- ? ~~N ≥ 30~~ NO
- or
- ? No STRONG Skew or outliers OK

**DO: If the conditions are met, perform the calculations.**

General Formula: Point Estimate ± Margin of Error

Specific Formula: $\bar{x}_{diff} \pm t^* \cdot \frac{s_{diff}}{\sqrt{n}}$

Work: $-.69 \pm 2.13 \cdot \frac{3.14}{\sqrt{16}}$

TABLE B

$-.69 \pm 1.67$

$(-2.36, .98)$

**CONCLUDE: Interpret your interval in the context of the problem.**

We are 95% confident that the interval from -2.36 to .98 captures the true mean difference in words remembered using re-reading and ~~quizzing~~ story telling.

$(-2.36 \quad .98)$

**CONCLUDE: Interpret your interval in the context of the problem.**

**6.** Do we have evidence that there is a difference in the average words remembered using rereading and story?

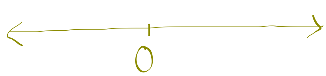
does the interval contain 0 ?
then NO

(+, +) re-reading better

(−, −) story strategy better

## Confidence Interval for Paired Data

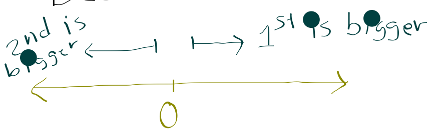
Important ideas:
Desribe Distribution



## Confidence Interval for Paired Data

Important ideas:

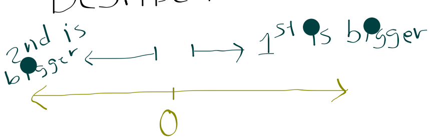
• Desribe Distribution




## Confidence Interval for Paired Data

Important ideas:

• Desribe Distribution




• Summary Stats: $\bar{x}_{diff}$ $s_{diff}$

## Confidence Interval for Paired Data

Important ideas:

Desribe Distribution

2nd is bigger ⟵ | ⟼ 1st is bigger

0

Summary Stats: $\bar{x}_{diff}$ $s_{diff}$

One-Sample t interval for $\mu_{diff}$

or paired t interval for a mean difference

## Confidence Interval for Paired Data

Important ideas:

Desribe Distribution

2nd is bigger ⟵ | ⟼ 1st is bigger

0

Summary Stats: $\bar{x}_{diff}$ $s_{diff}$

One-Sample t interval for $\mu_{diff}$

or paired t interval for a mean difference

$$\bar{x}_{diff} \pm t^* \cdot \frac{s_{diff}}{\sqrt{n}}$$

## Confidence Interval for Paired Data

Important ideas:

Describe Distribution

2nd is bigger ⟵ | ⟶ 1st is bigger

0

Summary Stats: $\bar{x}_{diff}$ $s_{diff}$

One-Sample t interval for $\mu_{diff}$

or paired t interval for a mean difference

$$\bar{x}_{diff} \pm t^* \cdot \frac{s_{diff}}{\sqrt{n}}$$

Calculator App

T Interval

---

### Conditions for Constructing a Confidence Interval About a Mean Difference

**Random:** Paired data come from a random sample from the population of interest or from a randomized experiment.

**10%:** When sampling without replacement, $n_{diff} < 0.10N_{diff}$.

**Normal/Large Sample:** The population distribution of differences (or the true distribution of differences in response to the treatments) is Normal or the number of differences in the sample is large ($n_{diff} \geq 30$). If the population (true) distribution of differences has unknown shape and the number of differences in the sample is less than 30, a graph of the sample differences shows no strong skewness or outliers.

Same as ch. 8 CI or $\mu$

FINAL EXAM REVISITED

Just do the "State" for now (because of time)

we'll just observe the others

**Final exam revisited** *Confidence interval for a mean difference*
In the preceding alternate example, a teacher used a matched pairs design to collect data on scores for two different versions of her final exam for 20 students. Recall that $\bar{x}_{diff} = \bar{x}_{A-B} = 3.2$ and $s_{diff} = 3.533$. Construct and interpret a 99% confidence interval for the true mean difference (Version A – Version B) in final exam scores for AP® Statistics students in this district.

**STATE** 99% CI for $\mu_{DIFF}$

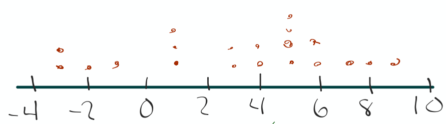
$\mu_{diff}$ = true mean difference (Version A - Vers. B) in the Final Exam Scores in the DISTRICT

**PLAN** One-Sample t interval for $\mu_{diff}$

Random: Rand. sample of 20 AP Students

10%: Assume 20 < 10% of all students in large district

Normal/Large Sample




Difference (A-B) in exam score

Dot plot doesn't show strong skewness or any outliers.

**DO** With 99% confidence and df = 20-1 = 19

$t^* = 2.861$

$$3.2 \pm 2.861 \cdot \frac{3.533}{\sqrt{20}}$$

$= (0.94, 5.46)$ ← Double Check w/ TInterval

**CONCLUDE** We are 99% confident that the interval from 0.94 to 5.46 captures the true mean difference (Version A - Ver. B) in final exam scores for AP Stats in this district.

Signif. tests for $\mu_{diff}$

Effects of Caffeine Withdrawal

$$\text{stand. test statistic} = \frac{\text{statistic} - \text{parameter}}{\text{std. dev. of stat}}$$

$$t = \frac{\bar{x}_{diff} - \mu_{diff}}{\frac{s_{diff}}{\sqrt{n_{diff}}}}$$

be sure to calculate the statistic.

## Significance Tests for Mean Difference ($\mu_{diff}$)

Researchers designed an experiment to study the effects of caffeine withdrawal. They recruited 11 volunteers who were diagnosed as being caffeine dependent to serve as subjects. Each subject was barred from coffee, colas, and other substances with caffeine for the duration of the experiment. During one 2-day period, subjects took capsules containing their normal caffeine intake. During another 2-day period, they took placebo capsules. The order in which subjects took caffeine and the placebo was randomized. At the end of each 2-day period, a test for depression was given to all 11 subjects. Researchers wanted to know whether being deprived of caffeine would lead to an increase in depression.

The table below contains data on the subjects' scores on the depression test. Higher scores show more symptoms of depression.

| Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Depression (caffeine) | 5 | 5 | 4 | 3 | 8 | 5 | 0 | 0 | 2 | 11 | 1 |
| Depression (placebo) | 16 | 23 | 5 | 7 | 14 | 24 | 6 | 3 | 15 | 12 | 0 |

Do the data provide convincing evidence at the $\alpha = 0.05$ significance level that caffeine withdrawal increases depression score, on average, for subjects like the ones in this experiment?

Consider using a histogram on GDC

State $H_0: \mu_{diff} = 0$ $H_a: \mu_{diff} > 0$

where $\mu_{diff}$ = true mean difference (Placebo − Caffeine) in depression test score for subjects like these.

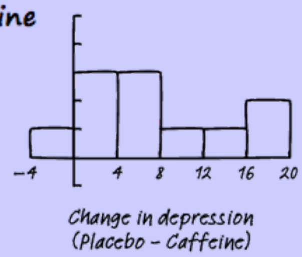
We'll use $\alpha = 0.05$

[ $H_a: \mu_{diff} < 0$
if using (Caffein − Placebo) ]

PLAN

**PLAN** Paired t test for $\mu_{diff}$

- Random: Researchers randomly assigned the treatments— placebo then caffeine, caffeine then placebo—to the subjects.✓
- Normal/Large Sample: The sample size is small, but the histogram of differences doesn't show any outliers or strong skewness. ✓




DO

$\bar{x}_{diff} = 7.364$ $S_{diff} = 6.918$

$n_{diff} = 11$

$$t = \frac{7.364 - 0}{\frac{6.918}{\sqrt{11}}} = 3.53$$

P-Value $df = 11 - 1 = 10$

$$\text{tcdf} = (3.53, 1000, 10) = .0027$$

(3.53 = lower, 1000 = upper, 10 = df)

CONCLUDE

Sice the P-Value of 0.0027 < $\alpha$ = .05 we reject Ho

We have convincing evidence that caffeine withdrawal increases depression test scores, on average, for subjects like the ones in the study.

Let's read bottom of p. 681

LCQ 8.2

↑ going to have to be a take home

10.3........75, 79, 85

+ LCQ 10.2