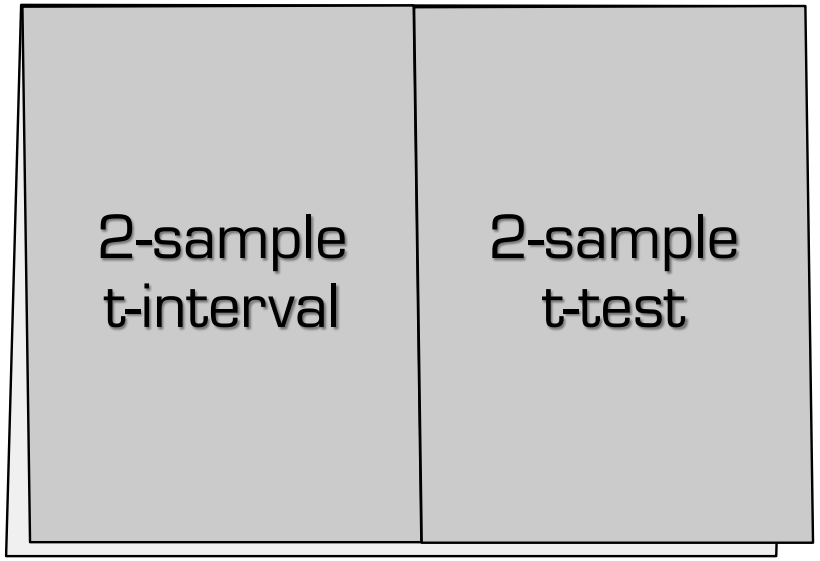
Inference with Means

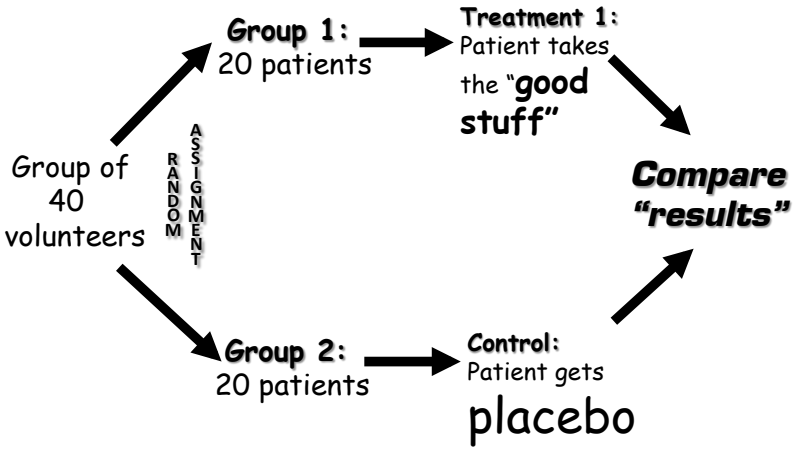
(two samples)

AP Statistics
Chapter 24

Update your foldables (outside, MIDDLE 2 flaps)



where today's 2-sample procedures fit it...



Do these situations involve *independent groups?*

1) Is there a difference in the mean amount of money spent by men and women for dinner each week? The mean amount of money spent on dinner for a random sample of 10 men and a separate random sample of 10 women were compared. None of the subjects were friends/relatives of one another.

Yes!

2) Is there a difference in the mean amount of money spent by men and women for dinner each week? 10 pairs of married men & women were surveyed, and the mean amounts of money spent on dinner were compared between the genders.

Nope!

Do these situations involve *independent groups?*

1) A group of 10 students were given the SAT before a special 6-week prep-class, and then took the SAT again after the class to determine the mean score increase for each student. **Nope!**

2) A group of 10 students at school "A" took the SAT, and another group of 10 students at school "B" took the SAT at about the same time. Schools "A" and "B" are in different cities. The scores between the two groups were measured to see if there is a significant difference between the scores from the two schools. **Yes!**

Update your foldables (inside, top half)

Define μ_1 and μ_2... *("true mean...")*

Conditions:

- ***Random Samples***
- ***Independent Groups***
- ***(10%)***
- ***Nearly Normal Condition***
  - *populations are normally distributed*
  - *n > 30 (for EACH group! not combined!)*
  - *Graph **BOTH** sets of data (outliers?)*

For a randomized experiment: *Check for RANDOM ASSIGNMENT of subjects to treatment groups*

df = (smaller n – 1) OR **use calculator**

Degrees of Freedom...

Approximation formula

$$df \approx \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)+\frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)}$$

Calculator does this automatically!

Update your foldables (inside, bottom half)

2-sample t-interval:

statistic ± crit. value × standard error

$$(\bar{x}_1-\bar{x}_2)\pm t^*\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

INTERPRET THE INTERVAL:
"We are ___% confident that the true difference in mean _____ is between _____ and _____."

INTERPRET THE CONF. LEVEL:
"If we repeat this method MAAANY times, about ___% of the resulting intervals will contain the true difference in the mean... _____."

2-sample t-test:

$H_O: \mu_1 - \mu_2 = 0$

$H_A: \mu_1 - \mu_2 \gtrless 0$

$$t = \frac{(\bar{x}_1-\bar{x}_2)-0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

AP Statistics – Inference with Two Sample Means

THE NOT-ENOUGH SHOES PROBLEM How many pairs of shoes do teenagers have? To find out, a group of AP Statistics students conducted a survey in which they selected two separate random samples of 12 male students and 12 female students from their school. Then they recorded the number of pairs of shoes that each respondent reported having. The data is displayed below.

Females:

12	13	15	15	19	21	22	24	26	31	34	41

Mean: 22.75 Standard Deviation: 8.9861 Number of students: 12

Males

4	5	5	6	7	8	10	10	11	12	14	17

Mean: 9.0833 Standard Deviation: 3.988 Number of students: 12

a) Construct and interpret a 95% confidence interval for the difference in the mean number of pairs of shoes owned between male and female students at this high school.

2-sample t-interval

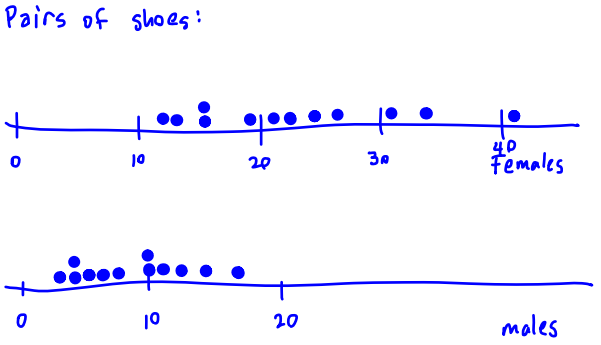
(use calculator!)

(7.6176, 19.7157)
df = 15.171

We are 95% confident that the true difference in the **mean** number of shoes owned between male and female students (female – male) at this high school is between 7.62 and 19.72 pairs of shoes.

Conditions for inference:

- **Independent Random Samples:** The data was collected via separate (thus, reasonably independent) random samples of male and female students.
- **Nearly Normal Condition:** The plots show some slight skewness, but with no major outliers, normality should be plausible for both groups.



b) Carefully interpret the meaning of the 95% confidence level in context.

If we repeated this method maaaaaaaaaaaaaany times, about 95% of the resulting intervals would contain the true difference in the mean number of shoes owned between male and female students at this high school

AP Statistics – Inference with Two Sample Means

"EACH DAY I AM GETTING BETTER IN MATH" A subliminal message is below our threshold of awareness but may nonetheless influence us. Can subliminal messages help students learn math? A group of 18 students who had failed the mathematics part of the City University of New York Skills Assessment Test agreed to participate in a study to find out. All received a daily subliminal message, flashed on a screen too rapidly to be consciously read. The 10 students in group "A" (assigned at random) were exposed to "Each day I am getting better in math." The control group of 8 students (group "B") was exposed to a neutral message, "People are walking on the street." All 18 students participated in the summer program designed to raise their math skills, and all took the assessment test again at the end of the program. The tables below gives data on the each subject's test score improvement:

Message "A" (10 students)

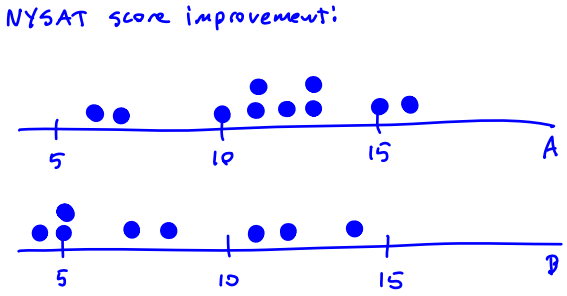
6	7	12	11	15	16	11	13	13	10

Mean: 11.4 Standard Deviation: 3.17

Message "B" (8 students)

11	5	4	8	14	5	7	12

Mean: 8.25 Standard Deviation: 3.69



At the 10% level of significance, do the data provide evidence that the mean NYSAT score improvement for students exposed to message "A" is higher than the mean NYSAT score improvement for students exposed to message "B"?

2-sample t-test

μ_1 = true mean score improvement for students receiving message "A"
μ_2 = true mean score improvement for students receiving message "B"

H_O: $\mu 1 = \mu 2$
H_A: $\mu 1 > \mu 2$

(use the calculator!)

t = 1.91356
p-value = 0.0382 **df = 13.9187**
α = 0.10

Since $p < \alpha$, we reject the Ho.
We have sufficient evidence that the mean score for students receiving subliminal message "A" is higher than for those receiving message "B".

Conditions:

- Students were randomly assigned to receive one of the two subliminal messages
- Based on graphs of the sample data for both groups, normality is plausible (no major outliers or obvious skew to either graph).