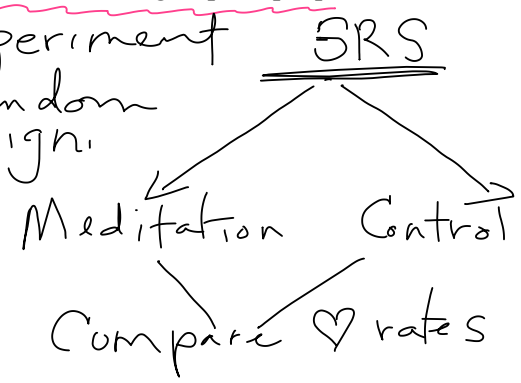
Math 247: Comparing Two Population Means (Section 9.5)

Measure stress level
cortisol
blood pressure } Quant.
♡ rate

Dependent Samples vs. Independent Samples

Suppose you're going to do your Senior Project on using meditation to reduce stress. How would you do this?

Using the same group of people ♡ rate	Using two different groups of people
Baseline ♡ rate Meditation "treatment" Measure ♡ rate again Before-and-After Study Design	Experiment SRS Random assign.  Compare ♡ rates
Paired t-Test: Used for these **Dependent** samples: • Before-and-After studies • Twin studies • Any situation where there is a natural pairing of the data values.	**Two Sample t-Test:** Used for **Independent** samples This is the most common situation in studies where you have two independent groups you're comparing, like the banded and unbanded penguins in Dr. Saraux's banding experiment.

Dependent Samples: Paired t-Test

SRS

Meditation and Resting Heart Rate: Suppose a random sample of 10 people is drawn to study the effect of meditation on resting heart rate. The entire group is assigned meditation for 20 minutes each day over a 6-week period. The raw data is given below.

Subject	1	2	3	4	5	6	7	8	9	10
Before (bpm)	74	67	83	70	77	83	70	75	72	88
After (bpm)	70	67	72	65	78	75	60	70	64	86
D = Difference	4	0	11	5	-1	8	10	5	8	2

Find the difference in heart rate for each subject.

If meditation didn't actually have a beneficial effect, we'd expect the difference in heart rate to be 0, on average.

is linked to a reduction in ♡ rate

Looking at this data, does it appear that meditation actually helped reduce heart rate?

Yes, most subjects had some reduction in resting ♡ rate

How can we tell if the reduction was *significant*?

Do Hypothesis Test ; Construct a CI
P-Value

INFERENTIAL STATISTICS

Here are the summary statistics for the two sets of data:		N	Mean	StDev
	Before	10	75.90	6.81
	After	10	70.70	7.56
	Difference	10	5.20	4.08

some variability WITHIN the group

less variability

Comment on the variability in the data for the Before and After groups vs. in the Difference data.

there is less variability in the differences BETWEEN than "groups" than there was WITHIN the groups

Before-and-After study design is useful because it reduces variability

Do our "signal" that something is going on is not drowned out by the "NOISE" of variability

Use StatCrunch to perform the computational step of a hypothesis test on the data **and** to find the Confidence Interval for the difference in means, using

(a) the Paired t-Test, — "Groups" are dependent — Correct

(b) the Two-Sample t-Test. — wrong test because it requires groups to be INDEPENDENT.

The data is on the wrightmath.info website. The results are below for reference.

Interpret the results of the hypothesis tests and the CI's.

(a) Paired T hypothesis test:

D = differences μ_D = mean of the differences

$\mu_D = \mu_1 - \mu_2$: Mean of the differences between Heartrate Before and Heartrate After

$H_0 : \mu_D = 0$ $H_A : \mu_D \neq 0$

$\mu_D = 0 \Rightarrow$ ZERO difference in Population

Hypothesis test results:

Difference	Mean	Std. Err.	DF	T-Stat	P-value
Heartrate Before - Heartrate After	5.2	1.289272	9	4.033284	0.003

Interpret Hypothesis Test: P-value = .003 < .05 = α

Reject H_0 and accept H_a

There was a significant difference in resting ♡ rate before and after the meditation program.

Paired T confidence interval:

$\mu_D = \mu_1 - \mu_2$: Mean of the difference between Heartrate Before and Heartrate After

95% confidence interval results:

Difference	Mean	Std. Err.	DF	L. Limit	U. Limit
Heartrate Before - Heartrate After	5.2	1.289272	9	2.2834642	8.1165358

Interpret CI: We are 95% confident that the difference in resting ♡ rate would, on average, between 2.3 bpm and 8.1 bpm if we applied this to the entire population. There is a significant reduction in ♡ rate, on average, between 2.3 bpm and 8.1 bpm.

Wrong test!

(b) Two sample T hypothesis test:

μ_1 : Mean of Heartrate Before μ_2 : Mean of Heartrate After

$\mu_1 - \mu_2$: Difference between two means

$H_0 : \mu_1 - \mu_2 = 0$ $H_A : \mu_1 - \mu_2 \neq 0$ (without pooled variances)

Hypothesis test results:

Difference	Sample Diff.	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$	5.2	3.216278	17.805913	1.6167756	0.1235

Interpret Hypothesis Test: P-value = .1235 > .05 = α

Fail to reject H_0!

The data does not provide evidence there is a significant difference in before-and-after groups ... Conflict? Signal is drowned out by the noise!

95% confidence interval results:

Difference	Sample Diff.	Std. Err.	DF	L. Limit	U. Limit
$\mu_1 - \mu_2$	5.2	3.216278	17.805913	-1.5624328	11.962433

Interpret CI: (-1.7, 11.96)

0 is captured by (within) the CI

⇒ ZERO difference is a possibility

so not a significant difference.

Follow-up: Comment on the importance of choosing the appropriate hypothesis test when analyzing data.

If you choose the wrong test, the results are not meaningful!

Independent Samples. If our two samples are not associated (i.e., they're not paired or dependent in some other way), then we use a Two Sample t-Test for Means (similar to the Two-Proportion z-test we used for proportions, back in Section 8.4.) - penguins

In comparing two populations, remember that individual values will vary and there may be overlap between the populations, but ON AVERAGE there can still be a significant difference between the populations.

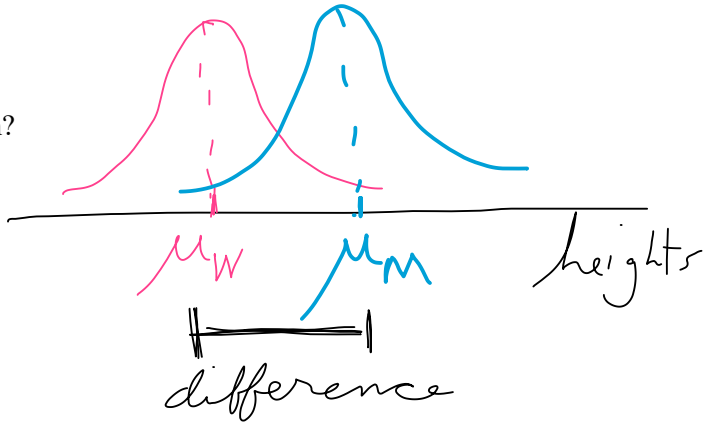
Consider heights of women and heights of men:

Are some individual women taller than some individual men?

Yes

On average, are men taller than women?

Yes



Two Sample t-test for Means (unpooled variance)

Hypotheses:

$H_0: \ \mu_1 - \mu_2 = 0$ There is ZERO difference, on average, between the two populations with regard to the variable of interest. 2 between means

$H_a: \ \mu_1 - \mu_2 \neq 0$ There IS a difference, on average, between the two populations with regard to the variable of interest.

(Note: The alternative hypothesis above is for a TWO-Tailed Test. We could also use a One-tailed test.)

Conditions:

1. **Random Samples and Independence within samples.** We have two random samples from the two populations. Each observation is independent from all others.
2. **Independence between samples.** The two samples themselves are independent from each other. The individuals in one sample are in no way associated with the individuals in the other sample.
3. **Sample Size Considerations.** Either the sample size in each sample is 25 or more ("large" samples) OR if, n < 25, ("small" samples) the underlying populations are each approximately normal.

(Central Limit Theorem) for the Difference of Means: Two Independent Samples If the conditions above are met, then the Sampling Distribution of the differences will follow the t-distribution*.

*Since we don't usually know the population standard deviation, the calculations involved in finding the Test Statistic will use the sample standard deviation. StatCrunch will do these calculations for us!

Standard Error (unpooled):

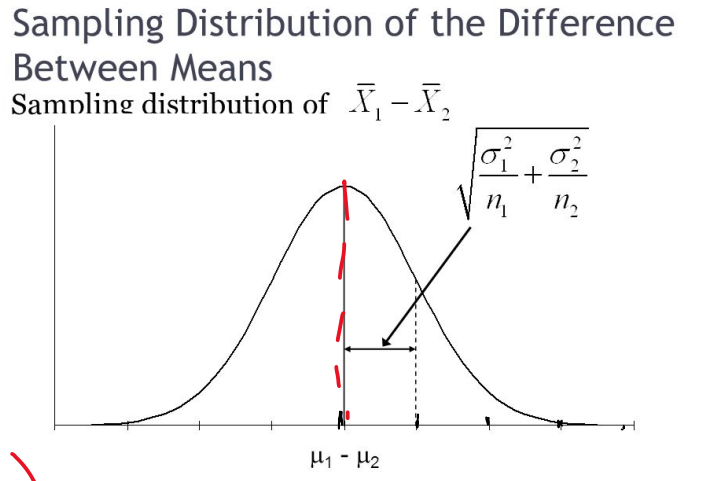
$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$SE_{est} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Test statistic: Use t-Stat

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE}$$

$df = ?$



F.Y.I.

By hand, we'd estimate degrees of freedom as df = least of $n_1 - 1$ and $n_2 - 1$

(This is just for reference…StatCrunch will do the compute step for us, and will use a more complex, and more accurate, way to find df.)

Example: Handedness and typing speed. Is handedness associated with how fast people type? Two random samples of left-handed students and of right-handed students who completed an online typing class are given a typing speed test (words per minute), and the results are compared. The 16 students in the right-handed sample had an average typing speed of 55.8 words per minute (wpm) with standard deviation of 5.7 wpm. The 9 students in the sample of left-handed students had an average speed of 59.3 wpm with a standard deviation of 4.3 wpm.

List and label the given values (Parking Lot!)

Sample 1: $n_1 = 16$, $\bar{x}_1 = 55.8$ wpm, $s_1 = 5.7$ wpm

Sample 2: $n_2 = 9$, $\bar{x}_2 = 59.3$ wpm, $s_2 = 4.3$ wpm

Looking at the sample data, does it appear that there's a difference in how fast people type, on average, based on whether they're left-handed or right-handed?

Yes, left-handers appear to be faster, on average

Is this difference *significant*? We don't know — the difference we see in the sample means may just be chance.

Conduct a hypothesis test to determine if there is a significant difference in typing speed in right-handed and left-handed people. Use a .05 level of significance.

μ = Mean typing speed for ALL people who take this typing course

1. $H_0: \mu_1 - \mu_2 = 0$ (Right − left) → ZERO difference between typing speeds, on average, for right and left handed people

$H_a: \mu_1 - \mu_2 \neq 0$ → There IS a difference

2. Prepare

Choose Two Sample t-Test for Means

1. Random sample? Yes, stated. Independence WITHIN samples? Assume
2. Independence BETWEEN samples? Yes, this is not a before-and-after or paired data situation
3. Large samples? NO, $n_1 = 16 < 25$ $n_2 = 9 < 25$ small samples. Assume typing speeds are normally distributed

3. Use StatCrunch for this step.

Hypothesis test results:

Difference	Sample Diff.	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$	-3.5	2.0211555	20.795052	-1.7316827	0.0981

$H_0: \mu_1 - \mu_2 = 0$

4. P-value = .0981 > .05 = α

Fail to reject H_0

There is not a statistically significant difference in typing speeds, on average, between right-handed and left-handed people, based on this data from this study