Today, we will:

- Do some review example problems – hypothesis testing (one sample)
- Review the pdf module: **Two Samples Hypothesis Testing** and do some examples

Example: Hypothesis testing

Given: A manufacturer claims that a plastic part is *at least* 6.00 cm long. You test the claim by performing a hypothesis test. You pick 30 parts at random from the assembly line, and carefully measure the length of each one. You calculate $\bar{x}$ = 6.053 cm and S = 0.104 cm.

To do: To what confidence level (%) can we claim that the manufacturer's claim is true?

Solution:

- This is a one-sided hypothesis test since we see "at least"
  (or one-tailed)
- Null hypothes → set μ_0 = manufacturers claim = 6.00 cm
  Set "side" to the least likely scenario
  Here, since our $\bar{x} > \mu_0$, the least likely scenario is that $\boxed{\mu < \mu_0}$ OUR NULL HYPOTHESIS

  [The alternative or research hypothesis is the opposite, i.e. that $\mu > \mu_0$]
- Calculate the critical t statistic: $t_{crit} = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}} = \dfrac{6.053 - 6.000}{0.104/\sqrt{30}}$

  $\boxed{t_{crit} = 2.79128}$
- Calculate the p-value @ $t = 2.79128$, $df = 30-1 = 29$, # tails = 1

  Table (use df = 19 table) $\boxed{p = 0.004595}$ (after interpolation)

  OR, Use EXCEL → = TDIST(2.79128, 29, 1) = $\boxed{p = 0.004595}$ ✓

  (t_{crit}, df, # tails)
- Interpret results:
  - We are 0.4595% confident that the null hyp. is true.

  ∴ We are 100 − 0.4595 = 99.5% confident that the manufacturer's claim is true, i.e. that $\mu > 6.053$ cm

  ACCEPT

• We would write this in a report as:

"We accept the manufacturer's claim to 99.5% confidence"

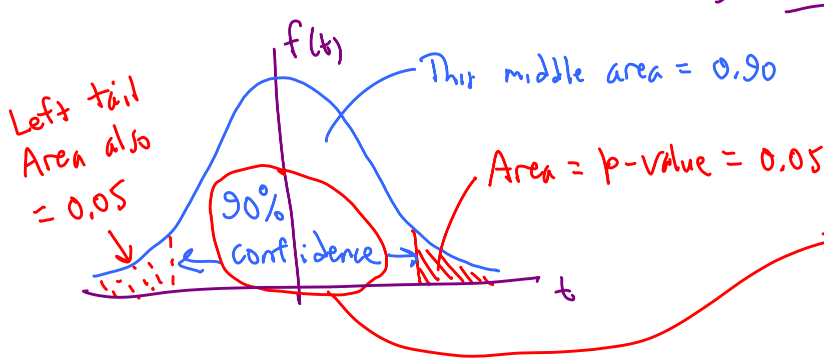
OR, "We accept the manufacturer's claim ($p = 0.0046$)"

ADDITIONAL QUESTION: WHAT MANUFACTURER'S CLAIM WOULD GIVE US EXACTLY 95% CONFIDENCE?

SOLUTION: METHOD A → TRIAL & ERROR → keep guessing μ_0 until $p = 0.0500$

METHOD B → DIRECT CALCULATION, WORKING BACKWARDS

• For 95% confidence with one tail, $\alpha/2 = 0.05$, & $p = 0.05$



★ BE CAREFUL OF FACTOR OF 2 ERRORS! Here, we look up $t_{crit} = t_{\alpha/2}$ in the 90% confidence level table, NOT 95% conf.

• Set $\alpha = 0.1$, $\alpha/2 = 0.05$, $c = 0.90$

Table of critical t values → @ 90% conf. & $df = 29$, get $t_{\alpha/2} = t_{crit} = 1.6991$

• Now, since we know t_{crit} or $t_{\alpha/2}$, we work "backwards":

$$t_{\alpha/2} = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \longrightarrow \mu_0 = \bar{x} - t_{\alpha/2}\frac{S}{\sqrt{n}}$$

$$= 6.053 - 1.6991\frac{0.104}{\sqrt{30}} = 6.0207$$

• BOTTOM LINE:

A manufacturer's claim of $\mu_0 = 6.02$ cm yields 95% confidence ★

OR, We are 95% confident that $\mu > 6.02$ cm

Example: Hypothesis testing

Given: We buy a gadget that is supposed to increase the gas mileage of our car. We take 6 trips *without* the gadget and 6 (nearly identical) trips *with* the gadget. The results:

x_A (mpg without gadget)	x_B (mpg with gadget)	$\delta = x_B - x_A$ (mpg)
25.6	26.2	0.6
27.3	27.1	−0.2
24.2	24.1	−0.1
28.7	29.2	0.5
23.6	24.5	0.9
25.1	24.9	−0.2

To do: Determine if there is a statistically significant improvement (increase) in gas mileage.

Solution: • This is a paired samples hypothesis test (n=6 for both A & B)

• Calculate sample means → $\boxed{\bar{x}_A = 25.75 \text{ mpg} \quad \bar{x}_B = 26.00 \text{ mpg}}$

Based on these averages, we would say there is an improvement. But is this improvement statistically significant?

• Calculate $\delta = x_B - x_A$ for each run (see table)

• Now use δ in our hypothesis test, as a one-tail hypothesis test, but with δ as our variable instead of x.

• Calculate $\boxed{\bar{\delta} = 0.25 \quad S_\delta = 0.4764}$ (Same as $\bar{x}_B - \bar{x}_A$ by the way, = sample mean of δ) (Sample standard deviation of the six δ values)

• Null hypothesis → $\mu_0 = 0$, & the "side" is set to the least likely scenario. Since our $\bar{\delta} > 0$ (improvement in mpg), we set null hyp. to $\boxed{\mu < \mu_0 \text{ or } \mu < 0 \text{ here}}$

Our null hypothesis is set to the least likely scenario, which is that gas mileage decreases rather than increases (thinking like a statistician)

• t statistic → $t = \dfrac{\bar{\delta} - \mu_0}{S_\delta / \sqrt{n}} = \dfrac{0.25 - 0}{0.4764/\sqrt{6}} = \boxed{1.285 = t \text{ statistic}}$

[Note, we use δ instead of x, but the procedure is otherwise the same as before]

- p-value → [We do not have a table for df = 5, so either look up in a table in a statistics book, or use Excel]

EXCEL → $p = \text{TDIST}(|t\text{-statistic}|, df, \#\text{ tails})$

[absolute value since t pdf is symmetric and Excel does not work if t is negative]

$= \text{TDIST}(1.285, 5, 1)$

We get $\underline{p\text{-value} = 0.1275} = \underline{12.75\%}$

- INTERPRETATION:
  - There is a 12.75% probability that the null hypothesis (the least likely scenario that mpg got worse) is true

OR
  - We are 87.25% confident that the gadget improves mpg

- CONCLUSION: Since 87.25% < 95%
  (i.e. since $p = 0.1275 > 0.05$),

  $\boxed{\text{We cannot accept the claim that this gadget improves gas mileage to at least 95\% confidence}}$

Bottom line — I would not waste my money on this gadget

NOTE: STATISTICALLY SPEAKING, WE CANNOT REJECT THE CLAIM, BUT WE ALSO CANNOT ACCEPT IT. WE SHOULD DO MORE TESTING

Example: Hypothesis testing

Given: [Continuation of previous example] We buy a gadget that is supposed to increase the gas mileage of our car. We take 6 trips *without* the gadget and 8 trips *with* the gadget. We do not attempt to pair up the tests. The results:

x_A (mpg without gadget)	x_B (mpg with gadget)
25.6	26.2
27.3	27.1
24.2	24.1
28.7	29.2
23.6	24.5
25.1	24.9
	26.5
	25.8

2 additional runs with the gadget

To do: Determine if there is a statistically significant improvement (increase) in gas mileage.

Solution: • Now, since $n_A \neq n_B$, this is not a paired-sample test

☆ $\boxed{\text{This is a hypothesis test with two independent samples}}$ ☆

Procedure:

• Calculate statistics → $n_A = 6$ $\quad n_B = 8$

$\bar{x}_A = 25.75$ $\quad \bar{x}_B = 26.0375$

$S_A = 1.9274$ $\quad S_B = 1.6405$

• Null hypothesis → Since $\bar{x}_B > \bar{x}_A$, we set the null hypothesis to the least likely scenario, i.e.,

$\boxed{\text{Null hypothesis is that } \bar{x}_B < \bar{x}_A}$

• Critical t statistic → $t = \dfrac{\bar{x}_A - \bar{x}_B}{\sqrt{\dfrac{S_A^2}{n_A} + \dfrac{S_B^2}{n_B}}} = -0.2941 = t$

• Use Welch's eq. to get df → $df = 10$

• Calculate p-value @ $|t| = 0.2941$, $df = 10$, # tails = 1 } in Excel, $p = \text{TDIST}(0.2941, 10, 1) = 0.3873$

· For $\boxed{p = 0.3873}$, we interpret:

- There is a 38.7% probability that the null hypothesis is true. In other words, there is a 38.7% probability that the gas mileage actually <u>decreases</u> due to this gadget.

<u>OR</u>

- We are 61.3% confident that the manufacturer's claim is true (i.e., that the gadget improves mpg)

<u>CONCLUSION</u> – Since 61.3% < 95%, we <u>cannot</u> <u>accept</u> the manufacturer's claim

[Again, we cannot <u>reject</u> the claim, but we <u>cannot</u> <u>accept</u> <u>it</u>]

In class, I will show how to do this problem in Excel, using the macro called

"t-test: Two Sample Assuming Unequal Variances"