Lecture 21 (Ch. 8)

In prev. example, we had $n = 64$, $\bar{x}_{obs} = 34.4$, $s_{obs} = 1.1$, and asked "Does data provide evidence to support $\mu > 34$?" Then

$H_0: \mu = 34$

$H_1: \mu > 34$

Note That H_1 plays an important role.

$\therefore$ p-value $= pr(\bar{x} > \bar{x}_{obs}) = prob(t > t_{obs}) = pr(t > 2.91) \overset{df=64-1}{=} 0.0025.$

$$t = \frac{\bar{x}_{obs} - \mu_0}{s/\sqrt{n}} = \frac{34.4 - 34}{1.1/\sqrt{64}} = 2.91$$

At $\alpha = 0.05$, p-value $< \alpha$. So, reject $H_0 (\mu = 34)$ in favor of $H_1 (\mu > 34)$

"In English:" There is evidence to support $\mu > 34$.

It is tempting to say The above "conclusion" (at $\alpha = .05$), That $\mu > 34$, is obvious and trivial, because after all, the observed sample mean $\bar{x}_{obs} = 34.4$ is greater than 34 already.

It's NOT obvious! Suppose The sample/data gave $\bar{x}_{obs} = 34.1$, ie. still larger Than 34 (μ_0). Then

$$t_{obs} = \frac{34.1 - 34}{1.1/\sqrt{64}} = .73 \Rightarrow \text{p-value} = prob(t > 0.73) = 0.24$$

This p-value is larger Than any reasonable α. So, we cannot reject H_0 in favor of H_1 even though The obs. sample mean <u>is</u> bigger than 34. 34.1 is larger than 34, but just not enough (in units of standard error, $\frac{s}{\sqrt{n}}$) to justify rejecting $H_0 (\mu < 34)$ in favor of $H_1 (\mu > 34)$.

"Rosetta Stone"

$\boxed{\alpha = .05}$ Data says: $n = 64$, $\bar{x} = 34.4$, $s = 1.1$

$\hookrightarrow t_{obs} = \dfrac{34.4 - 34}{1.1/\sqrt{64}} = 2.91$

Does data support $\mu > 34$? $\Longrightarrow$ "prior claim": $H_0: \mu < 34$

$H_0: \mu \le 34$ p-value $= \text{prob}(\bar{x} > \bar{x}_{obs} \mid \mu = 34) = \text{prob}(t > t_{obs})$

$H_1: \mu > 34$ $= \text{prob}(t > 2.91) = .0025 < \alpha$

$\therefore$ Reject H_0 ($\mu < 34$) in favor of H_1 ($\mu > 34$).

$\therefore$ Data does support $\mu > 34$.

Does data support $\mu < 34$?

$H_0: \mu \ge 34$ p-value $= \text{prob}(\bar{x} < \bar{x}_{obs} \mid \mu = 34) = \text{prob}(t < t_{obs})$

$H_1: \mu < 34$ $= \text{prob}(t < 2.91) = 1 - pr(t > 2.91) = 0.998 > \alpha$

$\therefore$ Cannot Reject H_0 ($\mu > 34$) in favor of H_1 ($\mu < 34$).

$\therefore$ Data does not support $\mu < 34$.

Does data contradict $\mu > 34$? $\leftarrow$ prior claim: $H_0: \mu > 34$

$H_0: \mu \ge 34$ p-value $= \text{prob}(\bar{x} < \bar{x}_{obs} \mid \mu = 34) = \text{prob}(t < t_{obs})$

$H_1: \mu < 34$ $= \text{prob}(t < 2.91) = 1 - pr(t > 2.91) = 0.998 > \alpha$

$\therefore$ Cannot Reject H_0 ($\mu > 34$) in favor of H_1 ($\mu < 34$).

$\therefore$ Data does not contradict $\mu > 34$.

Does data contradict $\mu < 34$?

$H_0: \mu \le 34$ p-value $= \text{prob}(\bar{x} > \bar{x}_{obs} \mid \mu = 34) = \text{prob}(t > t_{obs})$

$H_1: \mu > 34$ $= \text{prob}(t > 2.91) = .0025 < \alpha$

$\therefore$ Reject H_0 ($\mu < 34$) in favor of H_1 ($\mu > 34$).

$\therefore$ Data does contradict $\mu < 34$.

Note: small p-value $\Longrightarrow$ "does"
large p-value $\Longrightarrow$ "does not"

2-Sample Tests

We've done hyp. testing with p-values for testing a <u>single μ</u>:

$H_0: \mu = \mu_0 \qquad H_1: \mu \ \square\ \mu_0$

But, recall from the CI days, we also had 2 μ's. The hypotheses are then:

$H_0: \mu_2 = \mu_1 \qquad H_1: \mu_2 \ \square\ \mu_1$ (i.e. $\mu_2 - \mu_1 \ \square\ 0$)

It turns out we can solve a more general problem:

$H_0: \mu_2 - \mu_1 = \Delta \qquad H_1: \mu_2 - \mu_1 \ \square\ \Delta$

E.g. you may want to see if $\mu_2 - \mu_1$ exceeds 13, then $\Delta = 13$.
If you want to see if μ_2 and μ_1 are different, then $\Delta = 0$.
The point: Δ is determined by the question, not by data.
It's the 2-sample analog of μ_0.

⇒ Then If 2-samples are <u>indep.</u>, then assuming $H_0 = T$,

$$z = \frac{(\bar{x}_2 - \bar{x}_1) - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Recall $\begin{cases} z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \\ t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t,\ df = n-1 \end{cases}$

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t\text{-dist. with } df = \text{Welch.}$$

Then, p-values are computed just as before: See the "blue note"

$$\text{p-value} = \begin{cases} pr\left(\bar{x}_2 - \bar{x}_1 > (\bar{x}_2 - \bar{x}_1)_{obs} \mid \mu_2 - \mu_1 = \Delta\right) = pr(t > t_{obs}) & \text{if } H_1: \mu_2 - \mu_1 > \Delta \\ pr\left(\bar{x}_2 - \bar{x}_1 < (\bar{x}_2 - \bar{x}_1)_{obs} \mid \mu_2 - \mu_1 = \Delta\right) = pr(t < t_{obs}) & \text{if } H_1: \mu_2 - \mu_1 < \Delta \\ = \text{right area} + \text{left area} & \text{if } H_1: \mu_2 - \mu_1 \neq \Delta \end{cases}$$

⇒ If the two samples are <u>paired</u>: Make a new column:

x_1	x_2	$d = x_1 - x_2$
.	.	.
.	.	.
.	.	.
		$\bar{d}, s_d$

$$t = \frac{\bar{d} - \Delta}{s_d/\sqrt{n}} \sim t\text{-dist. } df = n-1$$

p-value computed as before.

Reconsider This past example. Let's do it with CI (now, with t) and with p-value.

Example: 82 students have picked-up their test, but 30 have not, even 1 week after the test was returned.
Call these 2 groups "Attenders" and "Non-attenders".

		n	$\bar{x}$	s
1	Non-attend	30	11.8	3.32
2	Attend	82	13.25	3.04

} sample

Important

μ_1 = mean of test1 for Non-attend students who have ever taken 390. population.

μ_2 = " " Attend students " " " " " .

Is there evidence from data that μ_1 and μ_2 are different?

We built the 2-sample (2-sided) 95% CI for $\mu_2-\mu_1$:

$$(13.25 - 11.8) \pm \overset{2.01}{\cancel{1.96}} \sqrt{\frac{(3.32)^2}{30} + \frac{(3.04)^2}{82}}$$

qt(.05/2, df=48) ← from Welch.

$$1.45 \pm \underset{2.01}{\cancel{1.96}}(0.693) = 1.45 \pm \underset{1.39}{\cancel{1.36}} = \cancel{(0.09, 2.81)} \; (0.06, 2.84)$$

∴ zero is not included in the CI ⇒ evidence that μ_1 and μ_2 are diff.

We also commented that because the whole CI is to the right of 0, there is evidence that $\mu_2 > \mu_1$.

⇒ Now, let's use the p-value method to answer the question "Is there evidence that $\mu_2 > \mu_1$?"

$H_0: \mu_2 - \mu_1 = 0$

$H_1: \mu_2 - \mu_1 > 0$

$$t_{obs} = \frac{1.45 - 0}{0.693} = 2.1 \text{ (not 2.01)}$$

$t^* \neq t_{obs}$

same conclusion

$$\text{p-value} = \text{prob}(t > 2.1 \underbrace{\mid \mu_2-\mu_1 = 0}_{\text{it's OK to not write this part.}}) \overset{\text{Table VI or } 1-pt(2.1,\,48)}{=} 0.0205$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{1}{n_1-1}\left[\frac{s_1^2}{n_1}\right]^2 + \frac{1}{n_2-1}\left[\frac{s_2^2}{n_2}\right]^2} = 47.91$$

Not important to remember this (Welch's) formula

At $\boxed{\alpha = .05}$, p-value < α.

∴ Reject H_0 in favor of H_1.

$\mu_2 < \mu_1$ $\mu_2 > \mu_1$

"In English": There is evidence that $\mu_2 > \mu_1$.

Do NOT Read This page until you are quite comfortable with CI's and p-values.

FYI

I have said repeatedly That $t^* \neq t_{obs}$.
And I stick by That!
But, it turns out There are some relationships between Them.
Here, is an illustration:

We have seen 2 ways of answering The question
Does data support $\mu_2 > \mu_1$?

1) Find 95% obs LCB for $\mu_2 - \mu_1$

$1.45 - \boxed{1.68}(.693) = 0.286$

$\therefore$ 95% confident That μ_2 exceeds μ_1 by at least 0.286

2) Test $H_0: \mu_2 - \mu_1 \le 0$ $H_1: \mu_2 - \mu_1 > 0$

p-value $= pr(t > \boxed{t_{obs}}) = pr(t > 2.1) = .0205$

$t_{obs} = \frac{1.45 - 0}{.693} = 2.1$

$\therefore$ At $\alpha = .05$, There is evidence for $\mu_2 - \mu_1 > 0$

But consider The following related questions:

3) At what conf. level is LCB_{obs} for $\mu_2 - \mu_1$ equal to zero?

$1.45 - t^*(.693) = 0 \Rightarrow \boxed{t^* = \frac{1.45}{.693} = 2.1}$

$\therefore$ Conf. level $= pr(t < t^*) = 1 - 0.0205 = .9795$

$\therefore$ 97.95% confident That μ_2 exceeds μ_1.

4) Does data support $\mu_2 - \mu_1 > 0.286$

$H_0: \mu_2 - \mu_1 \le 0.29$ $H_1: \mu_2 - \mu_1 > 0.286$

p-value $= pr(t > \boxed{t_{obs}}) = pr(t > 1.68) = 0.05$

$t_{obs} = \frac{1.45 - .286}{.693} = 1.68$

Note The relationships between t_{obs} & t^*.
But, again, until you are very comfortable with The 2 methods, keep t^* and t_{obs} separate.

Summary

We are done with 1-sample and 2-sample, z and t-tests, for paired and unpaired data, but all of that has dealt with pop. means.

Let's not forget pop. props. The procedure is the same

$z = \dfrac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0,1)$ but $\dfrac{p-\pi}{\sqrt{\frac{p(1-p)}{n}}}$ is NOT t!

1 sample

C.I. for μ_x:

$\bar{x} \pm z^* \dfrac{\sigma_x}{\sqrt{n}}$ | $\bar{x} \pm t^* \dfrac{s}{\sqrt{n}}$, $df = n-1$

C.I. for π_x:

$p \pm z^* \sqrt{\dfrac{p(1-p)}{n}}$

t-version does not exist

test for μ_x:

$H_0: \mu_x = \mu_0 \quad H_1: \mu_x \square \mu_0$

$z_{obs} = \dfrac{\bar{x}_{obs} - \mu_0}{\sigma_x / \sqrt{n}}$ | $t_{obs} = \dfrac{\bar{x}_{obs} - \mu_0}{s_x / \sqrt{n}}$, $df = n-1$

$p\text{-value} = \underset{1 \text{ or } 2}{\wedge} pr(\bar{x} \square \bar{x}_{obs})$

Test for π_x:

$H_0: \pi_x = \pi_0 \quad H_1: \pi_x \square \pi_0$

$z_{obs} = \dfrac{p_{obs} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$ ← No p! Because we assumed $H_0 = T$, i.e. $\pi_x = \pi_0$

$p\text{-value} = \underset{1 \text{ or } 2}{\wedge} pr(p \square p_{obs})$

2 sample

C.I. for $\mu_2 - \mu_1$:

$\bar{x}_2 - \bar{x}_1 \pm z^* \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ | $\pm t^* \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$, $df =$ Welch

(paired vs. unpaired)

C.I. for $\pi_2 - \pi_1$:

$p_2 - p_1 \pm z^* \sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$

Again No t!

Test for $\mu_2 - \mu_1$:

$H_0: \mu_2 - \mu_1 = \Delta \quad H_1: \cdots$

$z_{obs} = \dfrac{(\bar{x}_2 - \bar{x}_1)_{obs} - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ | $t_{obs} = \dfrac{(\bar{x}_2 - \bar{x}_1)_{obs} - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$, $df =$ Welch

$p\text{-value} = \underset{1 \text{ or } 2}{\wedge} pr(\bar{x}_2 - \bar{x}_1 \square (\bar{x}_2 - \bar{x}_1)_{obs})$

" $pr(t \square t_{obs})$

(paired vs. unpaired)

Test for $\pi_2 - \pi_1$:

$H_0: \pi_2 - \pi_1 = \Delta \quad \cdots$

$z_{obs} = \dfrac{(p_2 - p_1) - \Delta}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$

$p\text{-value} = \underset{1 \text{ or } 2}{\wedge} pr(p_2 - p_1 \square (p_2 - p_1)_{obs})$

$= pr(z \square z_{obs})$

hw_lect22_1

In hw_lect17_1 we used a CI to answer the question Does it appear that the true proportion of detective screws is not 2.5%. Here, answer the same question with the p-value approach. Specifically,

a) Set-up the appropriate hypotheses.

Hint: see summary Table

b) Compute the p_value (using the data in that hw)

c) At alpha =0.01, state the conclusion? Is it consistent with the conclusion from the CI approach?

hw_lect22_2

We are supposed, to transform our question into "Does data provide evidence for ...?" Usually the "..." is specified by you, the scientist. But just for practice, and to better understand the relationship between CIs and p-values, let's ask "Does data provide evidence that the difference (mu_2 - mu_1) is less than the upper side of the observed (2-sided) 2-sample CI?" Do not fix the Conf. level, i.e. don't pick a number for t*.

a) First, recall that the t* that appears in the formula for the CI is designed to satisfy pr(-t* < t < t*) = confidence level. Then, starting from that "self-evident fact," show that t* also satisfies pr(t< -t*) = (1-conf level)/2.

b) Setup H0, H1. Hint: look-up our formula for the observed 2-sample CI, and select the upper side. This step will also tell you what is delta for this problem.

c) Find the p-value in terms of the confidence level. Hint: Find the p-value, and then use part a.