STAT-1302; Lecture 9; Feb. 6, '24

Still in Section 10.2.

$H_0: \mu_1 = \mu_2$ where $\sigma_1 = \sigma_2 = \sigma$ but σ_1 and σ_2 are unknown.

$100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ when $\sigma_1 = \sigma_2$ but σ_i's are unknown:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Given

where $df = n_1 + n_2 - 2$.

See last lecture for S_p^2 formula.

Omit p-value approach.

A CI for $\mu_1 - \mu_2$ can be used to test $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$ in the usual way. That is,

i) Check to see if zero is in the CI.

ii) If "Yes" in part i), fail to reject H_0.

Otherwise, reject H_0.

Rule of thumb:

If $\dfrac{\text{larger } s}{\text{smaller } s} < 2$, then conclude $\sigma_1 = \sigma_2$.

Ex. A company claims that its medicine brand A, provides faster pain relief than brand B. A researcher tested both brands on two independent groups of randomly selected patients. Here are the results.

Brand	n_i	$\bar{x}_i$	s_i
A	25	$\bar{x}_1 = 44$	$s_1 = 11$
B	22	$\bar{x}_2 = 49$	$s_2 = 9$

Is the mean relief time from brand A less than the mean relief time for brand B? Let $\alpha = 0.01$. State any assumptions you are making.

Sol'n:

1. Parameter : $\mu_A - \mu_B$ where μ_A is the mean relief time due to brand A and μ_B is the mean relief time due to brand B.

2. $H_0 : \mu_A = \mu_B$

$H_1 : \mu_A < \mu_B$

σ_A and σ_B are unknown.

$\frac{11}{9} = 1.22 < 2 \Rightarrow$ Can assume $\sigma_1 = \sigma_2$

3. Test Stat.

$$t = \frac{(\bar{x}_A - \bar{x}_B) - 0}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_B}\right)}} = \frac{(44 - 49) - 0}{\sqrt{102.33\left(\frac{1}{25} + \frac{1}{22}\right)}}$$

$$= -1.69$$

$$S_p^2 = \frac{(n_A - 1) S_A^2 + (n_B - 1) S_B^2}{n_A + n_B - 2}$$

$$= \frac{(25-1)\,11^2 + (22-1)\,9^2}{25 + 22 - 2} = 102.33$$

$df = 25 + 22 - 2 = 45$

4. We reject H_0 if $t < -t_{0.01}$; $df = 45$.

i.e. if $t < -2.412$

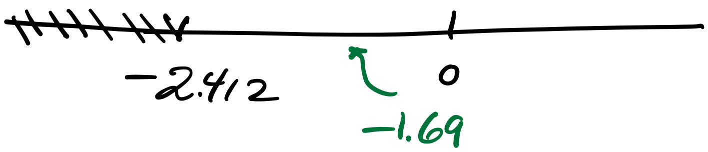


Table V

df	0.01
⋮	⋮
45	$\boxed{2.412}$

5. Decision: Since $t = -1.69 > -2.412$, we fail to reject H_0.

6. Conclusion: There is not enough evidence to suggest that mean relief time from brand A medicine is less than mean relief time

from brand B medicine.

Assumptions :

1. Two independent Samples.

2. Within each sample obs'ns are independent of each other. (Or write $\frac{n_i}{N_i} < 0.05$; $i=1,2$).

3. $\sigma_1 = \sigma_2$ and σ_i's unknown.

4. Since $n_1 = 25 < 30$ and $n_2 = 22 < 30$, relief times of both brands are normally distributed.

Ex. The following was obtained from two independent sample drawn from normal populations with unknown but equal standard deviations.

$n_1 = 21$, $\bar{x}_1 = 13.97$, $s_1 = 3.78$

$n_2 = 20$, $\bar{x}_2 = 15.55$, $s_2 = 3.26$

a) What is a point estimate of $\mu_1 - \mu_2$?

$$\bar{x}_1 - \bar{x}_2 = 13.97 - 15.55 = -1.58$$

b) Construct a 95% CI for $\mu_1 - \mu_2$.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \cdot S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\Leftrightarrow (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$df. = n_1 + n_2 - 2 = 21 + 20 - 2 = 39$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(21 - 1)\, 3.78^2 + (20 - 1)\, 3.26^2}{21 + 20 - 2}$$

$$= 12.50$$

$$\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{12.50\left(\frac{1}{21} + \frac{1}{20}\right)} = 1.10$$

$t_{\alpha/2} = ?$ $\quad n = 39$

$1-\alpha = 0.95$; $\alpha = 0.05$; $\alpha/2 = 0.025$

df	0.025
⋮	⋮
39	$\boxed{2.023}$

← Table V

CI : $-1.58 \pm 2.023 \times 1.10 = (-3.81,\ 0.65)$

We are 95% confident that $\mu_1 - \mu_2$ is between -3.81 and 0.65.

c) Test $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ at $\alpha = 0.05$.

The 95% CI for $\mu_1 - \mu_2$ is $(-3.81,\ 0.65)$.

Since $\mu_1 - \mu_2 = 0$ is in $(-3.81,\ 0.65)$ we fail to reject H_0.

§10.3 (FYI)

Set-up: Two independent Samples. Want to test $H_0: \mu_1 = \mu_2$; σ_1, σ_2 are unknown and unequal.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$df = ?$ (See §10.3).

§10.4 now.

Ex. A researcher wants to find the effect of a special diet on systolic blood pressure. She selected a random sample of seven adults and put them on this dietary plan for three months. The following table gives the systolic blood pressure before and after completion of the diet.

Subject	1	2	3	4	5	6	7
Before	210	180	195	220	231	199	224
After	193	186	186	223	220	183	233

Q'n: Was the diet effective in reducing systolic blood pressure? Let $\alpha =$.

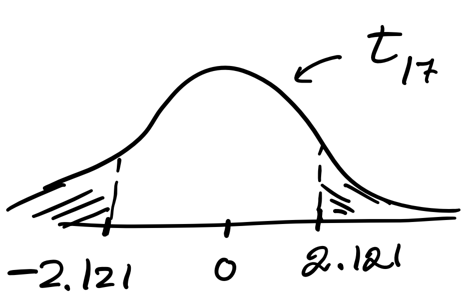
WileyPLUS : Ass't #2 ; Q'4 :

Ex. Test $H_0: \mu = 12.5$; $H_1: \mu \neq 12.5$; σ unknown

$n = 18$. ; $\bar{x} = 12.9$; $S = 0.80$. want: p-value.

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{12.9 - 12.5}{0.8/\sqrt{18}} = 2.121$$

$df = n - 1 = 18 - 1 = 17$

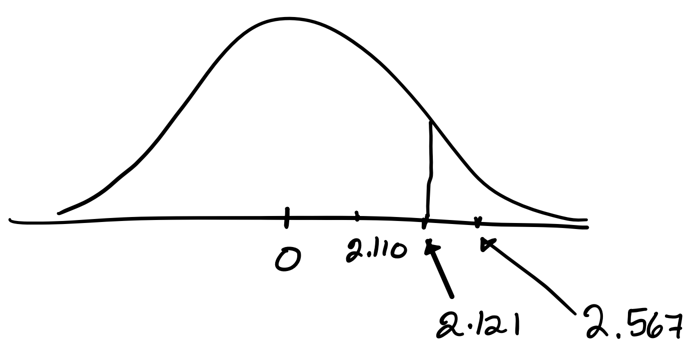


p-value $= 2 \times P(t > 2.121)$

However, the value of our test statistic $t = 2.121$ does not appear in the row where $df = 17$ (on Table V).

However, for $df = 17$, on that row we see that

$$2.110 < 2.121 < 2.567.$$



We want an approximation for $P(t > 2.121)$. On the one hand, it is larger that $P(t > 2.567)$. On the other hand, it is less than $P(t > 2.110)$.

From Table V,

df	0.025
⋮	⋮
17	2.110

and

df	0.01
17	2.567

So, $\quad 0.01 < P(t > 2.121) < 0.025$

$\therefore$ We double these bounds to get an approx'n for the p-value. That is,

$$\underbrace{2 \times 0.01}_{0.02} < P\text{-value} < \underbrace{2 \times 0.025}_{0.05}$$

Ex. Suppose we wish to test $H_0: \mu = 65$ vs. $H_1: \mu < 65$; σ is unknown.

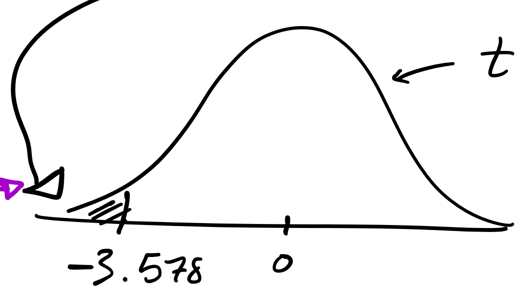
Given: $\bar{x} = 63.4$, $s = 3$ and $n = 45$.

What is the p-value?

test Stat: $\quad t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{63.4 - 65}{3/\sqrt{45}} = -3.578$

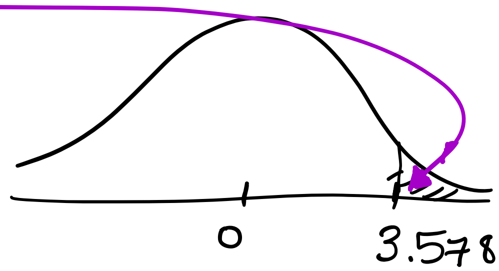
$df = n - 1 = 45 - 1 = 44.$

p-value is $P(t < -3.578)$



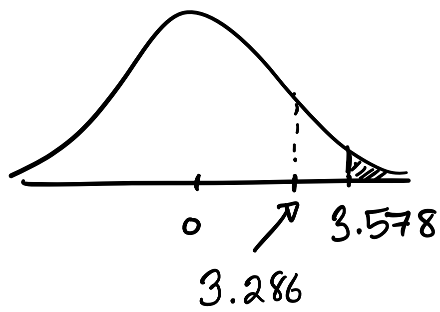
Same area

But, there are no negative numbers on the t-table (Table V). However, the t-distribution is symmetric about its mean (zero). So, we can find $P(t > 3.578)$



Now, there is no 3.578 on the row with $df = 44$ on Table V.

In fact, on that row, the largest number is 3.286.



From the figure, the right-tail probability of 3.578 is less than the right-tail probability of 3.286.

i.e. $\text{P-value} = P(t > 3.578) < \underbrace{P(t > 3.286)}_{0.001.}$