STAT_1302; Lecture 10; Feb. 8, '24.

Recall the very example from Lecture.

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
Difference (Before_After)	17 ✓	−6	9 ✓	−3	11 ✓	16 ✓	−9

Q'n: Was the diet effective in reducing systolic blood pressure? Let $\alpha = 0.025$.

* Can't use the two-sample t-test to Compare mean blood pressures before and after because the two Samples are not independent of each other.

* Before obs'ns and after obs'ns are dependent.

* But, the observations

$$d_i = \text{Before}_i - \text{After}_i \quad ; \; i = 1, \ldots, 7$$

do form an independent Set of obs'ns.

* Want to Compare μ_A and μ_B.

Jargon: The Two Samples in our example are said to matched Samples or matched pairs.

Observe, $d_i = \text{Before}_i - \text{After}_i$:

$$\bar{d} = \frac{\sum_{i=1}^{7} d_i}{7}$$

$$= \frac{\sum_{i=1}^{7} (\text{Before}_i - \text{After}_i)}{7}$$

$$= \frac{\sum_{i=1}^{7} \text{Before}_i}{7} - \frac{\sum_{i=1}^{7} \text{After}_i}{7}$$

$$= \bar{x}_{\text{Before}} - \bar{x}_{\text{After}}$$

$\bar{x}_{\text{Before}} - \bar{x}_{\text{After}}$ is a point estimator of $\mu_B - \mu_A$.

So is $\bar{d}$ an point estimator of $\mu_B - \mu_A$.

Message: To make inference about $\mu_B - \mu_A$ based on paired samples, use $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$ where d_i is the i-th paired difference.

Set-up: $d_1, \ldots, d_n$ is a random sample from a population with mean μ_d and unknown standard deviation σ_d. $\left(\frac{n}{N} < 0.05\right)$.

Hypotheses:

$H_0: \mu_d = 0$ vs. $H_1: \mu_d > 0$, $H_1: \mu_d < 0$
or $H_1: \mu_d \neq 0$.

Remark: Can also test $H_0: \mu_d = \mu_{d0}$ where μ_{d0} is not zero. (In this Course, $\mu_{d0} = 0$ always.)

Back to our example.

$H_0: \mu_d = 0$, $H_1: \mu_d > 0$.

Diet reduced B.P. means $d_i = \text{Before}_i - \text{After}_i$ are positive, on average.

$d_i = \text{Before}_i - \text{After}_i$.

Test Statistic:

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}}$$ Given

Under H_0, $t = \dfrac{\bar{d} - 0}{s_d/\sqrt{n}} \sim t_{n-1}$ where df = n − 1 with n = # of matched Pairs.

s_d = Sample Std. dev'n of the differences d_i.

$$s_d = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n} d_i^2 - \frac{(\sum d_i)^2}{n}\right]}$$ Given.

Test	Rejection Region
$H_0: \mu_d = 0,\ H_1: \mu_d > 0$	$t > t_\alpha$
$H_0: \mu_d = 0,\ H_1: \mu_d < 0$	$t < -t_\alpha$
* $H_0: \mu_d = 0,\ H_1: \mu_d \neq 0$	$\lvert t \rvert > t_{\alpha/2}$

df = n − 1 where n here is the number of matched pairs.

We can also test $H_0: \mu_d = 0$ vs. $H_1: \mu_d \neq 0$ using a $100(1-\alpha)\%$ CI for μ_d:

$$\bar{d} \pm t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}$$ where

d.f. $= n-1$; n = # of matched pairs.

Homework: A 95% CI for μ_d in our example is $(-4.98, 14.98)$. Use it to test $H_0: \mu_d = 0$ against $H_1: \mu_d \neq 0$.

Since zero is in the CI for μ_d, we fail to reject H_0. The result is not statistically significant at the 5% level.

Back to the B.P. example:

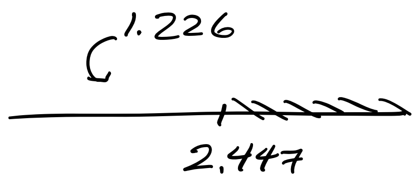
μ_d = mean of the difference between blood pressures ; d = Before − After

$H_0: \mu_d = 0$, $H_1: \mu_d > 0$. Let $\alpha = 0.025$.

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}} = \frac{5 - 0}{10.79/\sqrt{7}} = 1.226$$

Reject H_0 if $t > t_{0.025}$; $df = 7 - 1 = 6$.

From Table V, $t_{0.025;6} = 2.447$



Since $1.226 < 2.447$, we fail to reject H_0.

Conclusion: There is not enough evidence to suggest that the diet was effective in reducing Systolic B.P.

Next page for next example.

Ex. The manufacturer of a gasoline additive claims that the use of its additive increases gasoline mileage. A random sample of six cars was selected and these cars were driven for one week with the additive and one week without the additive. The following gives the miles per gallon with and without the additive.

	Car					
	1	2	3	4	5	6
Without	24.6	28.3	18.9	25.3	15.4	29.5
With	26.3	31.7	18.2	23.7	18.3	30.9
Difference (Without − With)	−1.7	−3.4	0.7	1.6	−2.9	−1.4

⤳ next page

Can you conclude that the use of the additive increases the gasoline mileage? Let $\alpha = 0.025$.

* The Samples are dependent because the with additive and without observations come from the same car.

Put $d_i = \text{Without}_i - \text{With}_i$; $i = 1, \dots, 6$

Assumptions: $n = 6 < 30$, we have to assume $d_1, \dots, d_6$ are a random sample from a normal population with mean μ_d and std. dev'n σ_d (unknown) ; $\frac{n}{N} < 0.05$.

implies $d_1, \dots, d_6$ are independent obs'ns.

$H_0: \mu_d = 0$

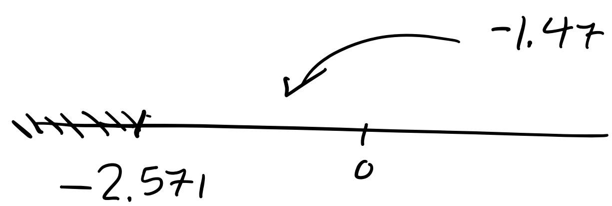
$H_1: \mu_d < 0$ ← additive increases mileage; $d_i = \text{Without}_i - \text{With}_i$.

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}} = \frac{-1.183 - 0}{1.97/\sqrt{6}} = -1.47$$

Reject H_0 if $t < -t_{0.025}$; $df = 6 - 1 = 5$.

From Table V, $t_{0.025;5} = 2.571$.

Reject H_0 if $t < -2.571$.



Since $-1.47 > -2.571$, we fail to reject H_0.

Conclusion: There is not enough evidence to suggest that the gasoline additive increases mileage.