```
Stat 145, Mon 25-Oct-2021 -- Mon 25-Oct-2021
Biostatistics
Spring 2021


-------------------------------
Monday, October 25th 2021
-------------------------------
Due))   PS10 due at 11 pm

-------------------------------
Monday, October 25th 2021
-------------------------------
Wk 9, Mo
Topic:: Student t distributions
Read::  Lock5 6.4-6.6, 6.13
HW::   WW ch06Part1 due Thurs.

Why t-distributions?
 - arise in context of quantitative data.  Why?

 - no need when doing one-proportion (nor two-proportion) inference

 - for 1-sample mean settings,  df = n-1

Practice with
 1. summarized data
     n=27 healthy white males, mean systolic bp: 114.9 with s=9.3
     (a) Give a 92% CI for true mean bp among healthy white males
     (b) Test mu = 118 vs. a 2-sided alternative

 2. raw data
     dogs, from boot package: use lvp variable (left ventricular pressure)
     BodyTemp50, from Lock5withR package
     t.test() command

 3. paired data
     hypotheses
     Wetsuits data
```

Test 2: Fri. Nov. 5

, n = sample size

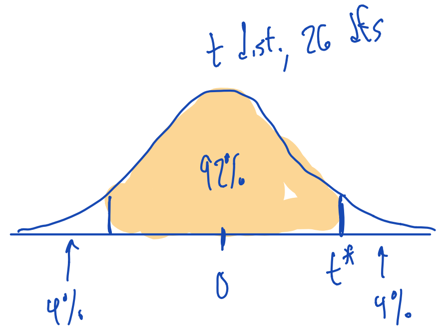
have raw data

Ex. 1) $\bar{x} = 114.9$, $s = 9.3$

(a) point est $\pm$ (critical value t^*)($SE_{\bar{x}}$)

$t^* = qt(0.96, df = 26) = 1.822$

$$114.9 \pm (1.822)\frac{9.3}{\sqrt{27}}$$



(b) $H_0: \mu = 118$ (μ_0 (null value)) vs. $H_a: \mu \neq 118$

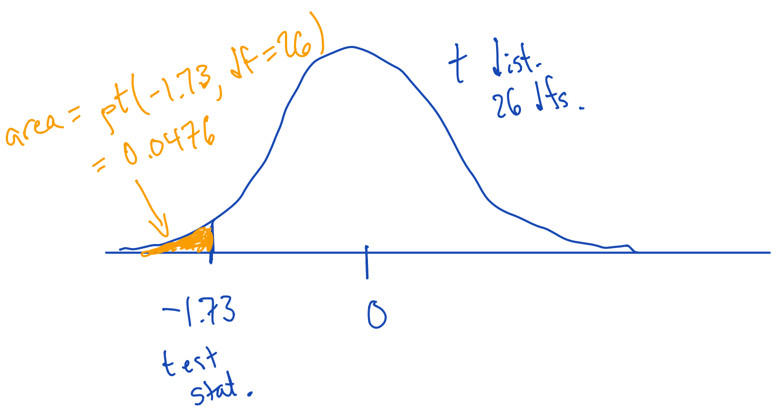
Guiding principle for conducting t-hypothesis tests: must standardize your point est.

Still have $\bar{x} = 114.9$

est. $SE_{\bar{x}} = 9.3/\sqrt{27}$

standardized score (test statistic)

$$t = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}} = \frac{114.9 - 118}{9.3/\sqrt{27}} = -1.732$$



P-value: $2(0.0476) = 0.0952$

Ex.) BodyTemp50 - BodyTemp variable

favstats says $\bar{x} = 98.26$, $s = 0.7653$

Now we could do a CI (like part (a) of last example)

If a 98% CI

$$SE_{\bar{x}} = 0.7653/\sqrt{50}$$

$$t^* = qt(0.99, df=49)$$

$$\bar{x} \pm (t^*)(SE_{\bar{x}})$$

Or, we could do hyp. test

$H_0: \mu = 98.6$ vs. $H_a: \mu \neq 98.6$

test statistic $t = \dfrac{\bar{x} - 98.6}{SE_{\bar{x}}}$

Convenient command in RStudio for doing all this: t.test( )

but you have to have raw data to use it.

(So, can't use it on Example 1.)

```
> t.test(~ BodyTemp, data= BodyTemp50)
```

or, for our needs

```
> t.test(~ BodyTemp, data= BodyTemp50, mu= 98.6, conf.level=0.98)
```

Ex.) Paired data

data set: Wetsuits

$H_0: \mu_{Diff} = 0$

From every swimmer/case - my interest is

not in wetsuit time

not in nowetsuit time

but in difference: wetsuit - nowetsuit

$n = 12$ here

You can see the 12 differences using

```
> with(Wetsuits, Wetsuit - NoWetsuit)
```

You can test H_0 vs. $H_a: \mu_{Diff} \neq 0$ using

```
> t.test(~ Wetsuit - NoWetsuit, data = Wetsuits)
```

(Wetsuit - NoWetsuit) indicates you want the two columns subtracted