24/5 AMs 5 Wednesday
Comparing 2 Samples
* H/W #5 due Fri. June 2nd (pg. 107-109)

C.S. 13 1/2 \( t \) Test  
Null Hyp.: \( M = 70 \) (recalibration not needed)
Alt. Hyp.: \( M \neq 70 \) (should recalibrate)

Hist. of \( \frac{\bar{y} - M}{s/\sqrt{n}} \) if null true

\[ t = \frac{77.8 - 70}{8.1/\sqrt{51}} = 2.16 \]

Evidence running against null but not very strongly Borderline Recalibrate

Why are there \((n-1)\) degrees of freedom w/ a sample of size \( n \)?

Whole premise: use \( \bar{y} \) to estimate \( M \);
this is like holding \( \bar{y} \) const. \( n=5 \) #s
but only \((n-1)=4 \) free to vary if the mean is fixed at 77.8

A pitfall of sig. testing:
(Statistical significance) \( \neq \) (Practical significance)

Ex.) New drug to combat high blood press.

\[
\begin{array}{c|c|c}
\text{After-Before} & \text{Blood Press} & \\
\hline
-3 mm & 8,000 & \\
-4 mm & 20 mm & \\
\end{array}
\]

Null Hyp.: Drug doesn't work: \( M = 0 \)
Alt. Hyp.: \( M \neq 0 \)
\( \hat{SE}(\bar{y}) = \frac{20 mm}{\sqrt{8,000}} \)
\( t = \frac{-1 mm - (0 mm)}{0.22 mm} = -4.5 \)
AMS 5  Wednesday

L-R Hist of \( \bar{y} \) if Null Hyp true

\[ SE = 0.22 \]

CLT

\[ \bar{y} = -1.0 \text{ mm} \]

\[ 0 \text{ to } 1.5 \]

\[ -4.5 \text{ to } 4.5 \]

\( \bar{y} \) is far less than 5% (1%), so this diff. is highly stat. sig., but a decline of only 1 mm on ave. is far too small to be practically sig.

So here (stat. sig: yes, prac. sig: no)

How did this happen?

Stat. sig. p/c \( \bar{y} \) is smll. \( \rightarrow \) \( \bar{y} \) is smll. b/c \( \bar{y} \) is big

\[ \bar{y} \text{ big} \Rightarrow \frac{-1}{0.22} \text{ fraction big when numerator is} \]

big (not true here! or when denominator smll. (true here))

\( \rightarrow \) Denom. smll. b/c \( SE \) smll. = \( \frac{A}{\sqrt{n}} \) = fraction smll. if numerator is smll. \( \frac{20 \text{ mm}}{\sqrt{8000}} \) (not true here) or denom. big (true here) Bad Exp. design b/c of too little data

Another exp. w/ too little data

\[
\begin{array}{c|c}
\text{after} & \text{before} \\
\hline
-25 & +5 \\
\hline
-15 & +5 \\
\hline
\end{array}
\]

\( \bar{y} = -10 \text{ mm} \)

\( A = 20 \)

\( SE(\bar{y}) = \frac{A}{\sqrt{n}} \)

\[ = \frac{20 \text{ mm}}{\sqrt{8}} = 7 \text{ mm} \]

95% C.I.: \(-10 \pm 1.4 \)

\( \bar{y} = \frac{-10 \text{ mm} - 10 \text{ mm}}{20 \text{ mm}} = -1.4 \frac{7 \text{ mm}}{} \)

L-R Hist of \( \bar{y} \) if null true

\[ \bar{y} \text{ true} \]

\[ SE(\bar{y}) = +1.4 \text{ mm} \]

\[ -1.4 \text{ to } 1.4 \]

\[ 1.4 \text{ to } 1.4 \]

\[ 0 \text{ to } 0 \]

This diff. is not stat. sig. (\( p = 16\% > 5\% \)) but it is prac. sig. (stat. sig: No, prac. sig: Yes)

Problem: Too Hi' Data
24/5 AMS 5  Wednesday

Fix: design exp. w/ right (5) so that stat. sig = pract. sig.

Paired

2 topics: (1) Paired comparisons; (2) analysis of 2 independent samples.

Paired comparisons:
outcome Sales Vol. (in cases) of product

Treatment: Discount vs. Standard

Marketing plan:
One possible design: get a bunch of stores (e.g. n = 120), randomize 100 to ① & 20 to ②
design 1

Better design possible?

PCFs: (Don't have to worry about PCFs in design 1 as far as validity of exp. is concerned b/c randomization will make ①, ② groups similar on PCFs) but can we increase efficiency (accuracy) of exp.? A: Yes! think about PCFs @ design time.

Main PCF: Overall sales Vol. on all products: How can we improve design by using this PCF?

Design #2

PCF
High
Low
med

Basic idea: hold PCF const. in creating matched pairs of stores

 assigns ① & ② in pairs at random
24/5  AMS 5  Wednesday

Design 2 is also valid but will be more efficient
b/c PCF was controlled for in design.

Paired comparisons: Design #2 ... ?

NOTE:

If you mess with one of the paired columns it would
destroy the design setup.

If they're independent it won't matter

Paired Difference  (1-sample) — know how to
(2-sample)  analyze this