19/5 AMS 5 Friday
Significance Testing

C.S. #12 Flex Time (significance testing)

Pop. All employees under flex time

Sample The observed employees

1. D.S. Possible Vs

N = 2,000

# Days absent

$\text{mean } \mu = 6.3$

$\text{SD } \sigma = ?$

H0: Null hypothesis ($H_0$): (skeptic's position: Nothing new)

$\text{null } \mu = 6.3$ (days)

$\text{Flex time doesn't work}$

Alternative hypothesis ($H_a$): (something new)

$\text{Flex time works}$

$\text{alternative } \mu < 6.3$ (days)

Long run mean

$\text{E}_\text{null}(\bar{Y}) = 6.3 = \mu_0$

Long run mean

$\text{SD}_\text{null}(\bar{Y}) = \frac{5.4}{\sqrt{100}}$

Use: $\text{SE}_\text{null}(\bar{Y}) = \frac{5.4}{\sqrt{100}}$

Long run hist. (pg. 2)

One approach to inference about $\mu$: CI.s Another: significance tests

1. Null Hypothesis ($H_0$): (skeptic's position: Nothing new)

$\text{null } \mu = 6.3$ (days)

$\text{Flex time doesn't work}$

Alternative Hypothesis ($H_a$): (something new)

$\text{Flex time works}$

$\text{alternative } \mu < 6.3$ (days)

2. Try the null on for size—see if the discrepancy between (how data came out) vs. (how data should have come out if null was true)

* If the discrepancy = null not good
  if large favor alternative
  if not large favor null
Distance Measure (Test Statistic)

\[
\frac{(\bar{y} - \mu_0)}{\sigma/\sqrt{n}} = \frac{\text{Obs. mean} - 6.3}{\text{SE}(\bar{y}) \text{ if null true}} = \frac{-0.9}{0.3} = -3 = Z
\]

Called a Z test!!

* Long run Hist. of \( \bar{y} \) if null true

\[\text{SE} = 0.3 \text{ days} \]

\[5.4 - 6.3 = \text{signal} \]
\[0.3 = \text{noise} \]
\[-3 = s \text{.u.} \]

\[P\text{-value} = \text{Chance if null is true, of getting data as x-treme as, or more extreme than, what we got.} \]

Here \( P = 0.15\% \) !!

Final form of reasoning: If \( P \) small = favor alt. !!
\nIf \( P \) big = favor null !!

How small is small? Conventional answer: \( P \leq 5\% \)
then the diff. is statistically significant (Stat. Sig.)

\( P \leq 1\% \) is highly Stat. Sig.

In this ex. \( P = 0.15\% \), Diff. between 6.3 (null) & 5.60 (Data) is highly Stat. Sig., so conclusion: on the face of it, flex. time seems to work.

So we favor the alternative
14/5 AM 5 Friday

95% C.I.: \( \bar{y} \pm 2 \times \text{SE}(\bar{y}) \)

\[
5.4 \pm 2 (0.3) =
\]

\[
(5.4 \pm 0.6 \text{ days})
\]

Alt. way to judge if a diff. is Stat. sig. is: Compute 95% C.I. & see if "null" value of \( \bar{y} \) is in the C.I. or not; if it's not, Diff. is stat. sig.

C.S. #13 (pg. 103)

46% vs. 33% is large in practical terms (i.e. this diff. is practically significant (pract. sig.))? But is it Stat. sig.? (i.e., could this diff. be explained by unlucky sampling?)

Null Hyp. (\( H_0 \)): His sampling method is like SRS

Alt. Hyp. (\( H_a \)): No it's not like SRS