11/8/18
Wave
Moretto

Lecture #13: Inference for Proportions

inferential summary

pop.	unknown pop quantity of interest	p = pop % of lab animals that would turn left (food)
sample	estimate	$\hat{p} = 83\%$
imag. data	give or take for $\hat{p}$ as est. of p	$\hat{SE}(\hat{p}) = 11\%$
	95% CI for p	$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (68\%, 100\%)$

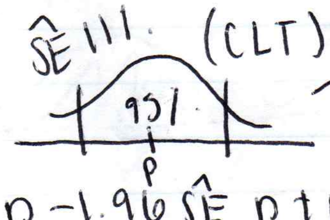
① EV of $\hat{p} = E_{IID}(\hat{p}) = E_{IID}(\bar{y}) = \mu = p$, so $E_{IID}(\hat{p}) = p$

② SE of $\hat{p} = SE_{IID}(\hat{p}) = SE_{IID}(\bar{y}) = \frac{\sigma}{\sqrt{n}}$

- Reminder: if a pop. has only 2 values in it, pop SD $= \sigma =$ $\left(\text{larger value}\right) - \left(\text{smaller value}\right) \cdot \sqrt{\left(\text{proportion larger values}\right) \cdot \left(\text{proportion smaller values}\right)}$
  - here, larger value = 1, smaller value = 0 so pop SD $= \sigma = \cancel{(1-0)}\sqrt{p(1-p)}$
- with 0/1 pop $\sigma = \sqrt{p(1-p)}$ new formula

$\Rightarrow$ so $SE(\hat{p}) = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$ so $\boxed{SE_{IID}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}}$

- so $\hat{SE}_{IID}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.837)(0.17)}{12}}$
  $= 0.108 = \boxed{11\%}$ ↑ don't use %, must be decimal

③ 

long run hist. of $\hat{p}$ using empirical rule

- There are methods for building CIs for building CIs for pop. proportions w/ small n - no time to look at, so just use large-sample approximation and pretend n is big enough to use CLT

- so approx 95% CI for p is $\hat{p} \pm 1.96\,\hat{SE}(\hat{p})$

$$= \hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- Here 83% ± 1.96 (11%)

61% 83% 100% 105%

truncate

embarassing

- when n is large

$\hat{p} \pm 1.96\,\hat{SE}_{IFD}(\hat{p})$

Devil's Advocate

50% 61% 83% 100%

50% devil's advocate

- since 50% is not in the 95% CI, the diff. btwn 50% and 83% is stating (is probably real), we don't agree w/ devil's advocate theory

- CIs are only 1 way to do a stastical inference - hypothesis tests (Neyman + pearson) and significance tests (Fisher)
  - these aren't as good as CI

null hypothesis (Ho)	$\mu_0 \rightarrow \mu = 24.3°C$	theory correct ①
alternative hypothesis HA	$\mu \neq 24.3°C$ 2-sided alternative	theory wrong ②

95% CI

24.4 25.0 25.6°C

24.3 theory probably wrong

① The diff between $\mu_0 = 24.3°C$ and $\bar{y} = 25.0°C$ is due to unlucky random sampling (this is a logical possibility)

② No, diff btwn 24.3 and 25 is real

• Neyman's logic: try null on for size; see if discrepancy between:

(how data came out) (*) vs. (how data should have come out if null true) (**)

is large;
- if yes, favor alt ("reject null")
- if not, favor null ("fail to reject null")

(*) (how data came out): $\bar{y} = 25.0$
vs. (**) (how data should have come out if null true)

$E_{FFD}(\bar{y})$ (if null true) $= \mu_0$

Measure discrepancy?

$$\boxed{\frac{\bar{y} - \mu_0}{(\text{SE of } \bar{y} \text{ if null is true})}} = \frac{25.0°C - 24.3°C}{0.27°C}$$

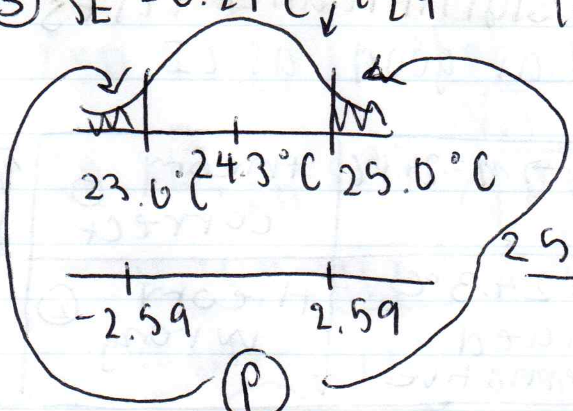
t test

$$= \frac{+0.7°C}{0.27°C} = \frac{\text{"signal"}}{\text{"noise"}} = +2.59 = t$$

"the t statistic"

③ $\hat{SE} = 0.27°C$, t_{24}

long run hist of $\bar{y}$ well true, accounting for uncertainty in σ



$$\frac{25.0°C - 24.3°C}{0.27°C} = 2.59 = t$$

• Chance, if null true, of getting data as extreme as, or more extreme than, what you got = numerical surprise measure = p value

- p-value is small favor alt, if p is big, favor null

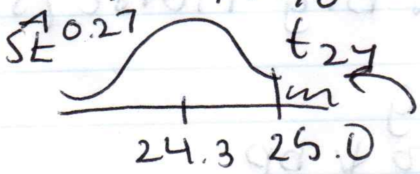
Q: How small is small enough for a p-value in order to "reject null" and favor alt?

A: Answering this question carefully is hard work and ppl are lazy, so people ~~worry~~ mostly appeal to conventional ideal

- Math fact: when hyp. testing is done w/ 2-sided alt. its conclusion is identical to that of CI approach

What can wrong w/ hypothesis testing

① (2 sides alt $\mu \neq \mu_0$) ⟷ (2 tailed t test = 2 tailed p value) → 1.6%



1-tailed $p = \frac{1.6\%}{2} = 0.8\%$

(1-side alt) ⟷ (1-tailed test)

- null: my theory wrong
- alt: my theory right
- Here I want to reject null ⟷ I want a small p-value, some journals so rigid about ~~stat sig~~ statsig

-p-hunting → torture data until p comes out small

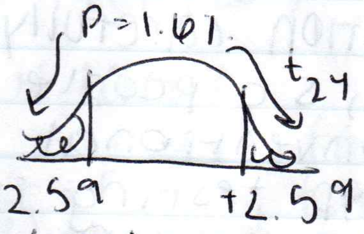
- (some) Journals will only accept paper if $p \leq 5\%$.

→ suppose do a test w/ my data get a 2-tailed value 8% (from 2-side alt) → can't publish. But quick way to ↓ p-value prent real alt. is 1-sided (4%), can publish

② why CIs better than p-values: crab data

null: $\mu = 24.3°C$ 2 tailed $p = 1.6\%$

alt: $\mu \neq 24.3°C$ $(n = 25)$



• If all you know is this, can't even tell if $\bar{y}$ came out above or below μ_0

• $t = \pm 2.59 = \boxed{\dfrac{\bar{y} - \mu_0}{s/\sqrt{n}}} = \dfrac{\bar{y} - 24.3}{s/\sqrt{25}}$

$= \boxed{\dfrac{\text{signal}}{\text{noise}}}$ → no way to tell how big signal $(\bar{y} - \mu_0)$ is or how big noise $(s/\sqrt{n})$ is

③ statsig ≠ practsig

ex) Have new drug to ↓ bp