May 10 (1)

K5

May 10, 2005 Lecture Notes   Significance Testing

Pitfalls for inference:

(1) How good is SE (in general) as give or take for $\bar{y}$ or $\hat{p}$?

obs_1 = truth + bias + random $error_1$

⋮

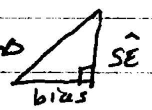
do over + over again obs_n = truth + bias + random $error_n$

= mean (average) $\bar{y}$ or $\hat{p}$ = truth + bias + (average n random errors)

$\hat{SE}$ measures the likely size of this

($\bar{y}$ or $\hat{p}$ − truth) = bias + (average of n random errors)

If no bias, $\hat{SE}$ = likely amount by which $\bar{y}$ or $\hat{p}$ will differ from the truth.

If bias, (math fact)  (like the hypotenuse)

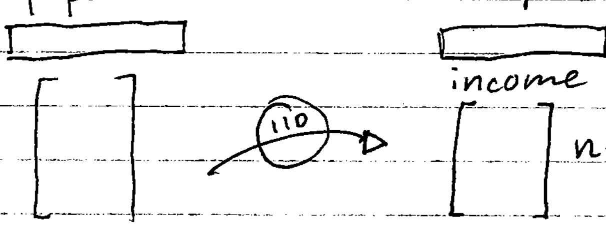
$\left(\begin{array}{l}\text{likely amount by} \\ \text{which } \bar{y} \text{ or } \hat{p} \\ \text{will differ from} \\ \text{the truth}\end{array}\right) = \sqrt{(\text{bias})^2 + (\hat{SE})^2}$ } can be much bigger than $\hat{SE}$

If no bias, reduces $\hat{SE}$ (which we already know to be the case)

punchline: If (substantial) bias is present in sampling method, the $\hat{SE}$ is likely to (substantially) understate your real margin of error (HW4 #2)

(2) How Inference can fail completely: (trying to find average income of people's parents who are in AMS 5)

ex)



mean μ = ?
SD σ = ?

mean $\bar{y}$ = \$60,000
SD s = \$30,000

95% CI for μ:

$\bar{y} \pm 2\hat{SE}_{NO}(\bar{y})$ =

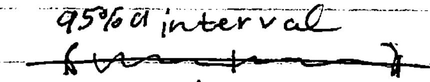
$\bar{y} \pm 2\frac{s}{\sqrt{n}}$ =

$\$60,000 \pm 2\left(\frac{\$30,000}{\sqrt{140}}\right)$ =

$\$60,000 \pm 2(2535) \approx$

$\$60,000 \pm \$5,000$

95% CI interval



1. 95% CI for what? for μ
2. what is μ? population mean
3. What's the population + do we actually have a random sample from it?

<u>possible population</u>:

(P1) everyone enrolled in AMS 5 this quarter who came to class today (If this is the population of interest, we have 0 uncertainty about μ + the right answer for μ is $\bar{y} = \$60,000 \pm \$0$)

In this case, sample = pop + there is no uncertainty left in pop

(P2) All students enrolled at UCSC in Spring 2005 (This pop is okay and sample ≠ pop, but class might be a <u>biased</u> sample from this pop, + the size + direction of this bias is hard to pin down)

- So $\$60,000 \pm \$5,000$ is probably too narrow to reflect our real uncertainty about μ of (P2), etc.
- <u>punchline</u>: If you can't identify an uncertainty population from which your data are arguably <u>like</u> a random sample, then don't do inference (CI or significance tests)

<u>2 ways inference can fail</u>: (HW4 #2)

(1) <u>conceptual failures</u>: whole idea is silly b/c sample is not related randomly to the right population

(2) <u>technical failures</u>: ex: if the data shows a time trend

Population (conceptual): $\boxed{\text{all possible measurements}}$

$N: \infty$ $\begin{bmatrix} \vdots \end{bmatrix}$

mean μ = ? = truth

SD σ = ?

$\xrightarrow{\text{like IID (no bias)}}$

sample: $\boxed{\text{observed measurements}}$

$\begin{bmatrix} \vdots \end{bmatrix}$ n =

mean $\bar{y}$ =

SD S =

How can you verify a model with a fictitious population?
- can't verify totally. A good check: make time trend plot of observations.

(see next page for graphs)