AMS 5

This time: Estimating a population Percentage

Next time: significance testing

readings: FPP ch. 26

5-5-05 ①

DM

Dara Morton

2 Survival Skills Concerning Statistical Models:

① which data set (pop., Sample, I.D.S.) is being referred to?

② Is the question about ~~individu~~ individual elements of the data set or a summary?

Finishing Case Study 10 Example (#1 in HW4)

(1.) About 68% of way bills in the sample were in the range $28.09 ± $0.69.

- This is False because the question is refering to the individual elements of the sample.
- make it true by replacing $0.69 with the S.D. of the sample, $31.40.

(2) About 68% of the waybills in the population were in the range of $28.09 ± $0.69

- False because elements in population data set differ from μ by amount roughly σ in size, and our best guess of σ is $s = \$31.40$

Case Study 11

Only 617 out of 950 students can add 1/2 + 1/3

- (oh my!)

• Devil's Advocate: "I think that the real % of all 17 yr. olds who can add fractions is 90%, and you just got a really different answer in your sample by unlucky sampling!"

- Plausible arguement in need of rebutting

• This is a statistical problem because we are reasoning from the sample to the population.

- the sample is known, but not the pop.

yes = 1
no = 0

Statistical Model
(filled in as we went along)

pop.
all 17 yr. olds in U.S. in 1975

can add fractions?

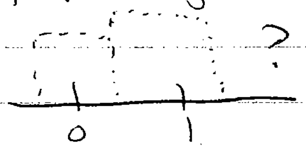
N = Big [1, 0, 1, 0, ..., 1]

(like SRS, like IID) →

$\mu = p = ?$

SD = $\sigma = \sqrt{p(1-p)}$

pop histogram:



Sample
observed 17 yr olds in US in 1975

can add fractions?

[1, 0, ..., 0] n = 950

$\bar{y} = \hat{p} = \frac{617}{950} = 65\%$

S = 617

(hypothetical IID)

hypothetical Sample

[⋮] n = 950

$\hat{p} = ?$ (ex 62%)

actual sample histogram

I. D. S
all possible $\hat{p}$'s

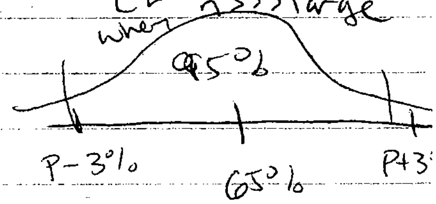
[65%, 62%, ⋮]

longrun mean = $E(\hat{p}) = p$

longrun SD = $\hat{SE}(\hat{p}) = 1.5\%$

long run hist.

CLT applies when n >> large



Standard Error $SE_{IID}(\hat{p})$

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

. we dont know p, but we can guess using $\hat{p}$

Estimate:

$$\hat{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.5\%$$

$$\sqrt{\frac{(.65)(.35)}{950}} \times 100\%$$

− $SE(\hat{p}) = \frac{\sigma}{\sqrt{n}}$

math fact

$\sigma = \sqrt{p(1-p)}$ in a pop w/ 1s ones and zeros

Inferential Summary

unknown quantity of interest	p = pop % of all US 17 yr olds who can add fractions
estimate	$\hat{p} = 65\%$
give/take *	$\hat{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.5\%$
95% CI *	$\hat{p} \pm 2\hat{SE} = 65\% \pm 3\% = (62\%, 68\%)$