this time: large-sample frequentist statistical inference

Ams 132

19 Jan 17

next time: Bayesian story

read ① DS ch. 7

quality of care:

outcomes of care: ex. mortality; full recovery from pneumonia

processes of care: → ex. how often were vital signs measured?

Ams 131:

$E_{RS, IID}(\hat{\theta}) = \theta$

standard error

$$SD_{RS,IID}(\hat{\theta}) = SE_{RS,IID}(\hat{\theta})$$

mean

$$= \frac{\sigma}{\sqrt{n}}$$

math fact (131):

pop. SD σ of binary (1s, 0s)

pop with mean θ =

$$\sigma = \sqrt{\theta(1-\theta)}$$

if $I_i = \begin{cases} 1 & \text{with prob. } \theta \\ 0 & 1-\theta \end{cases}$

then $I_i \sim \text{Bernoulli}(\theta)$ &

$$E_{RS}(I_i) = \theta$$

$$V_{RS}(I_i) = \theta(1-\theta)$$

$$SD_{RS}(I_i) = \sqrt{V_{RS}(I_i)} = \sqrt{\theta(1-\theta)}$$

so $SD_{RS,IID}(\hat{\theta}) = SE_{RS,IID}(\hat{\theta})$

$$= \sqrt{\frac{\theta(1-\theta)}{n}}$$

uncertainty : $n \uparrow$ $SE(\hat{\theta}) \downarrow$

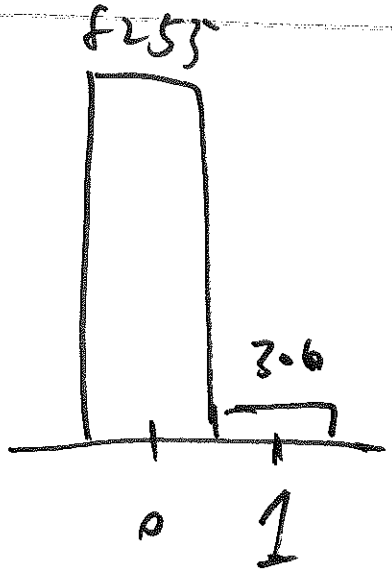
Central Limit Theorem (CLT):

As long as n is large, the RS dist. of the random variable $\hat{\theta} = \bar{Y}$ (sample mean based on a sample of size n) will look a lot like the normal curve

the closer the pop. dist. is to normal to begin with, the smaller n needs to be;

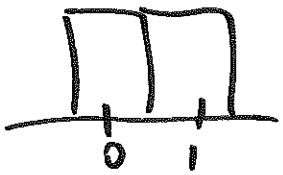
& if the pop. dist. <u>is</u> normal then the RS dist. of $\bar{Y} = \hat{\theta}$ <u>is</u> normal for all n

pop. dist. here



badly non-normal:

n = 112 may not be big enough



closest to normal