7.2 Sample Proportions

(mostly for categorical variables)

Q: How well does $\hat{p}$ estimate p?
statistic — parameter

A: Look at the sampling distribution (SOCS)

center: The mean of the sampling distribution of $\hat{p}$ is: $\mu_{\hat{p}} = p$

spread: The standard deviation of the sampling distribution of $\hat{p}$ is:

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

n = sample size

as n increases σ will decrease

* only if our sample is less than 10% of the population

shape: As n increases, the sampling distribution of $\hat{p}$ become more "normal"

if $n \cdot p \geq 10$
and $n(1-p) \geq 10$
then we can perform Normal calculations

Ex: Suppose that 80% of high school students in CSD are planning to attend a 4-year college. What is the probability that an SRS of size 125 will give a result within 7 percentage points of the true value?

p — n

$P(.73 \leq \hat{p} \leq .87)$

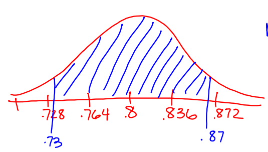
$\mu_{\hat{p}} = .8$

There are more than 1250 H.S. students in CSD $\rightarrow$ $\sigma_{\hat{p}} = \sqrt{\frac{.8(1-.8)}{125}} = .036$

is it Approx. Normal? yes.

$n(p) = 125(.8) = 100 \geq 10$ ✓

$n(1-p) = 125(.2) = 25 \geq 10$ ✓



normalcdf(.73, .87, .8, .036)
lower bnd, upper bnd, mean, st. dev.

$= .948$

There is about a 95% chance that our SRS (n=125) will give a $\hat{p}$ within 7% of the true proportion.

* see p. 438