7.2 Sample proportions

Is $\hat{p}$ a good estimator of p?

$\hat{p} = \frac{3}{10} = .3$

Takeaways:

Center: $\mu_{\hat{p}}$ was pretty close to p since $\hat{p}$ is an unbiased estimator.

Spread: $\sigma_{\hat{p}}$ gets smaller as n gets larger

Shape: sampling distribution of $\hat{p}$ was approximately Normal, depending on n + p

What is the connection X and $\hat{p}$

Let X = number of orange candies

$$\hat{p} = \frac{\text{\# of orange candies}}{\text{sample size}} = \frac{X}{n} = \frac{1}{n}X$$

If X is a binomial random variable:

$$\mu_X = np = \mu_{\hat{p}} = \frac{1}{n}np = p$$

$$\boxed{\mu_{\hat{p}} = p}$$

$$\sigma_X = \sqrt{np(1-p)}$$

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

$$= \frac{1}{\sqrt{n^2}}\sqrt{np(1-p)}$$

$$= \sqrt{\frac{np(1-p)}{n^2}}$$

$$\boxed{\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}}$$ make sure 10% condition is met

- As n increases, the sampling distribution of $\hat{p}$ becomes approximately normal as long as the <u>Large counts condition</u> is met.

Large counts condition:

$np \geq 10$ <u>and</u> $n(1-p) \geq 10$

Ex: check your understanding on p.445

1. $\mu_{\hat{p}} = p = .75$

2. $\sigma_{\hat{p}} = \sqrt{\frac{.75(1-.75)}{1000}}$

$= .0137$

$n \le \frac{1}{10} N$

1000 all 18-29 year old internet.

$\boxed{\text{There are definitely more than 10,000 internet users between 18\&29.}}$

3. $np \ge 10$ $n(1-p) \ge 10$

$1000(.75)$ $1000(.25) \ge 10$

$750 \ge 10$ $250 \ge 10$

Yes, you use the normal approx.

Using the Normal Approximation for $\hat{p}$

I. Nference based on a population proportion p is based the sampling proportion $\hat{p}$.

Ex: A superintendent of a large school district wants to know the proportion of students that plan on attending a 4 yr. university. Suppose 80% of all students are planning on attending a 4 year university. What is the probability that an SRS of size 125 will give a result within 7% points of its true value.

1. $\mu_{\hat{p}} = .80$

2. $\sigma_{\hat{p}} = \sqrt{\frac{.80(1-.80)}{125}}$

   $\sigma_{\hat{p}} = .0358$

$n \le \frac{1}{10}N$

125

A large school district has more than 1250 kids, so 10% condition is met

3. Normal?

   $125(.8) \ge 10$ $\quad$ $125(.2) \ge 10$

   $100 \ge 10$ $\quad$ $25 \ge 10$

   yes we can use the normal approximation

4 Normal Calculation

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

$z \le \frac{.73-.8}{.0358}$ $\quad$ $z \le \frac{.87-.8}{.0358}$

$z \le -1.96$ $\quad$ $z \le 1.96$

$.025$ $\quad$ $.975$

$.975 - .025 = .95$

There is a 95% chance the actual score of a 125 size sample will be with 7% points.