Section 5.3: cont'd

Tuesday, February 27, 2018 3:32 PM

so we are estimating the proportion of people/objects in a population who have a specified characteristic

let x = number in sample with characteristic
n = sample size

$$\hat{p} = \frac{x}{n} = \text{sample proportion}$$

from the binomial, x is a random variable with
mean $\mu = np$
$\sigma = \sqrt{npq}$

then $\hat{p}$ is also a random variable with mean

mean p

and $\sigma = \frac{\sqrt{npq}}{n} = \sqrt{\frac{pq}{n}}$

so the average value of $\hat{p}$ is p and the uncertainty in p is called the standard error

$$SE = \sqrt{\frac{pq}{n}}$$

one last thing:

if n is large and p is neither too close to zero nor too close to one, then we can approximate the binomial with a normal curve

specifically, if $np > 5^*$ and $nq > 5$ } must test both!

we can assume that $\hat{p}$ will be normally distributed

(why? idea is that $\mu \pm 2\sigma$ lies within the range $[0, n]$)

*note: we are using 5 (five) in this course
- some textbooks use 4
10
6
8