CHAPTER 7: SAMPLING DISTRIBUTIONS

Part 2: Sample Proportions

Launch w/ activity @ rossmanchance.com

[redacted]

How good is a statistic $\hat{p}$ as an estimate of the parameter p?

What are the effects of n and p on shape, center, and spread?

Ask yourself: What would happen if we took many many samples? The sampling distribution answers this questions.

Review Sampling Reese's Pieces – describe the shape, center, spread of sampling distribution

Shape: roughly symmetric, probs normal
Center: mean of 400 sample is .449 which is pretty dang close to p=0.45
Spread: SD is .105 (use own samples here)

As n increases, spread decreases, does not change center
As p decreases, the graph just shifts (center changes), does not change spread

[redacted]

The sample proportion of successes is closely related to X:

We know the mean and standard deviation of a binomial random variable:

$\hat{p}$= count of successes/size of sample = X/n

Mu = np and sd = sqrt(np(1-p))

Because $\hat{p}$= X/n = (1/n)X, we're just multiplying the random variable X by a constant (1/n) to get the random variable $\hat{p}$.
Recall from chapter 6 that multiplying by a constant multiplies both the mean and sd

Mu = 1/n(np) = p (confirming that $\hat{p}$ is an unbiased estimator of p)
Sd = 1/nsqrt(np(1-p)) = sqrt((np(1-p)/n^2) = sqrt(p(1-p)/n) (as sample size increases, spread decreases

Remember that when np and n(1-p) are at least 10, a Normal distribution can be used to approximate the sampling distribution of $\hat{p}$

[redacted]

Choose an SRS of size n from a population of size N with proportion p of success. Let $\hat{p}$ be the sample proportion of successes. Then:

1.

2.

3.

1. The mean of the sampling distribution of $\hat{p}$ is mu$\hat{p}$ = p
2. The standard deviation of the sampling distribution of $\hat{p}$ is sd$\hat{p}$ = sqrt((p(1-p)/n) as long as the 10% condition is satisfied n$\leq$ 1/10(N)
3. As n increases, the sampling distribution of $\hat{p}$ becomes approximately normal. Before you perform Normal calculations, check that the large counts condition is satisfied: np$\geq$10 and n(1-p)$\geq$10

**these formulas are given on the ap exam sheet
**Gotta check these conditions, otherwise normal calculations will be flawed!

[redacted]

About 75% of young adult Internet users (age 18 to 29) watch online videos. Suppose that a sample survey contacts an SRS of 1000 young adult Internet users and calculates the proportion $\hat{p}$ in this sample who watch online videos.

1. What is the mean of the sampling distribution of $\hat{p}$? Explain.

2. Find the standard deviation of the sampling distribution of $\hat{p}$ Check that the 10% condition is met.

3. Is the sampling distribution of $\hat{p}$ approximately Normal? Check the Large Counts condition is met.

4. If the sample size were 9000 rather than 1000, how would this change the sampling distribution of $\hat{p}$?

1. 0.75
2. .0137, there are more than 10(1000) =10,000 young adult internet users, so10% condition met
3. Yes, both np=750 and n(1-p)=250 are at least 10
4. Still be normal with mean 0.75, however standard deviation would be smaller at .0046

[redacted]

A polling organization asks an SRS of 1500 first-year college students how far away their home is. Suppose that 35% of all first-year students attend college within 50 miles of home.

Problem: Find the probability that the random sample of 1500 students will give a result within 2 percentage points of this true value. Show your work.

Break down: we want to know if this particular sample of 1500 students will give us something close to 35% (within 2% anyway)