

CHAPTER 7: SAMPLING DISTRIBUTIONS

Part 2: Sample Proportions

Launch w/ activity @ [rossmanchance.com](https://www.rossmanchance.com)



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



The sample proportion of successes is closely related to X :

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

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

$$\text{Mu} = np \text{ and } \text{sd} = \sqrt{np(1-p)}$$

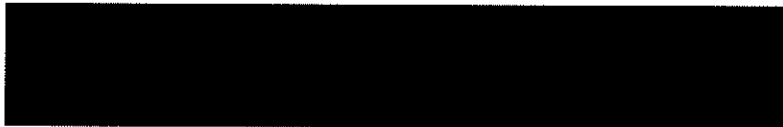
Because $\hat{p} = X/n = (1/n)X$, we're just multiplying the random variable X by a constant $(1/n)$ to get the random variable \hat{p} .

Recall from chapter 6 that multiplying by a constant multiplies both the mean and sd

$$\text{Mu} = 1/n(np) = p \text{ (confirming that } \hat{p} \text{ is an unbiased estimator of } p)$$

$$\text{Sd} = 1/n\sqrt{np(1-p)} = \sqrt{(np(1-p)/n^2)} = \sqrt{p(1-p)/n} \text{ (as sample size increases, spread decreases)}$$

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}



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!



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)