Chapter 7: Sampling Distributions

Section 7.2
Sample Proportions

The Practice of Statistics, 4th edition – For AP*

STARNES, YATES, MOORE

The Sampling Distribution of $\hat{p}$

What did you notice about the shape, center, and spread of different sampling distribution?

Shape : If we have enough samples(see criteria) the sampling distribution of $\hat{p}$ can be approximated by a Normal curve.

Center : The mean of the distribution is $\mu_{\hat{p}} = p$. This makes sense because the sample proportion $\hat{p}$ is an unbiased estimator of p.

Spread : For a specific value of p, the standard deviation $\sigma_{\hat{p}}$ gets smaller as n gets larger.

The Sampling Distribution of $\hat{p}$

In Chapter 6, we learned that the mean and standard deviation of a binomial random variable *X* are

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

Since $\hat{p} = X/n = (1/n)\cdot X$, we are just multiplying the random variable X by a constant $(1/n)$ to get the random variable $\hat{p}$. Therefore,

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

$\hat{p}$ is an unbiased estimator of p

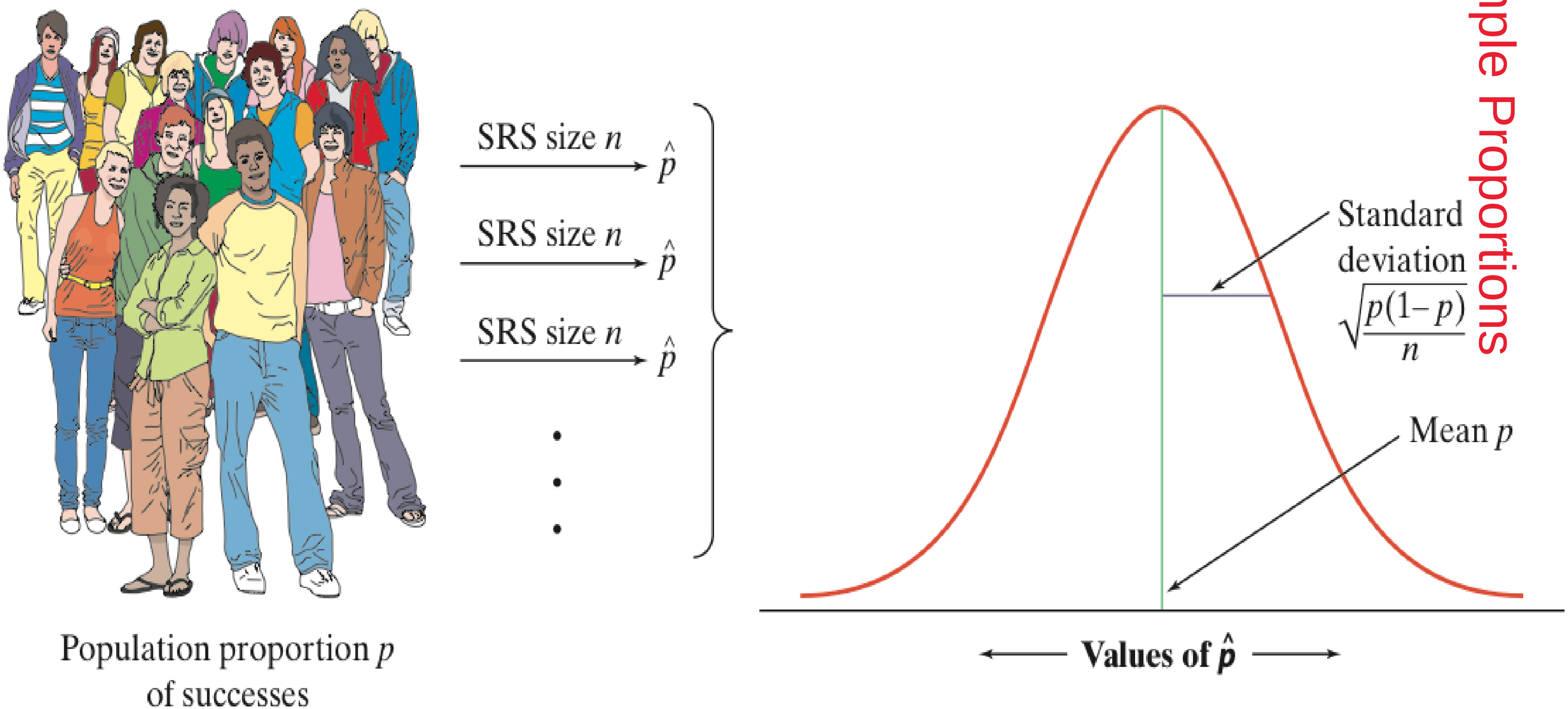
$$\sigma_{\hat{p}} = \frac{1}{n}\sqrt{np(1-p)} = \sqrt{\frac{np(1-p)}{n^2}} = \sqrt{\frac{p(1-p)}{n}}$$

As sample size increases, the spread decreases.

Sample Proportions

The Sampling Distribution of $\hat{p}$

We can summarize the facts about the sampling distribution of $\hat{p}$ as follows:



When the sample size is large enough for np and $n(1-p)$ to both be at least 10 (the Normal condition), the sampling distribution of $\hat{p}$ is approximately Normal.

The Sampling Distribution of $\hat{p}$

Sampling Distribution of a Sample Proportion

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

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

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

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

as long as the *10% condition* is satisfied: $n \leq (1/10)N$.

As n increases, the sampling distribution becomes **approximately Normal.** Before you perform Normal calculations, check that the *Normal condition* is satisfied:

$np \geq 10$ and $n(1 - p) \geq 10$.

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

STEP 4

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 actually attend college within 50 miles of home. What is the probability that the random sample of 1500 students will give a result within 2 percentage points of this true value?

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

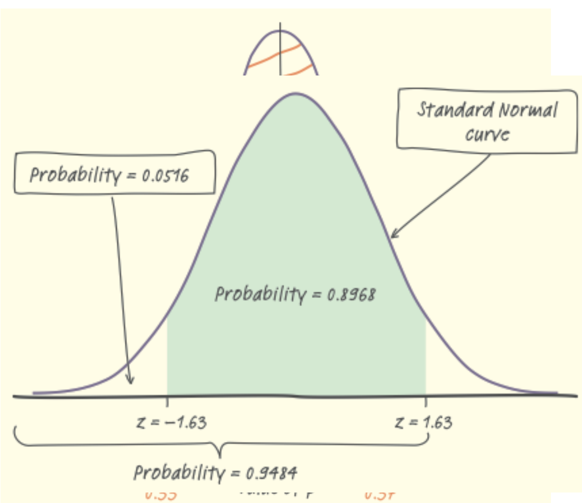
Inference about a population proportion p is based on the sampling distribution of $\hat{p}$. When the sample size is large enough for np and $n(1-p)$ to both be at least 10 (the Normal condition), the sampling distribution of $\hat{p}$ is approximately Normal.

STEP 4

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 actually attend college within 50 miles of home. What is the probability that the random sample of 1500 students will give a result within 2 percentage points of this true value?

STATE: We want to find the probability that the sample proportion falls between 0.33 and 0.37 (within 2 percentage points, or 0.02, of 0.35).

PLAN: We have an SRS of size *n* = 1500 drawn from a population in which the proportion *p* = 0.35 attend college within 50 miles of home.



$$\mu_{\hat{p}} = 0.35 \qquad \sigma_{\hat{p}} = \sqrt{\frac{(0.35)(0.65)}{1500}} = 0.0123$$

DO: Since np = 1500(0.35) = 525 and $n(1-p)$ = 1500(0.65)=975 are both greater than 10, we'll standardize and then use Table A to find the desired probability.

$$z = \frac{0.33 - 0.35}{0.123} = -1.63 \qquad z = \frac{0.37 - 0.35}{0.123} = 1.63$$

$$P(0.33 \le \hat{p} \le 0.37) = P(-1.63 \le Z \le 1.63) = 0.9484 - 0.0516 = 0.8968$$

CONCLUDE: About 90% of all SRSs of size 1500 will give a result within 2 percentage points of the truth about the population.