Welcome to Trimester 2

Pick Up the Warm Up

A **statistic** is a number that describes some characteristic of a sample.
A **parameter** is a number that describes some characteristic of the population.

Sample statistic		Population parameter
$\bar{x}$ (the sample mean)	estimates	μ (the population mean)
$\hat{p}$ (the sample proportion)	estimates	p (the population proportion)
s_x (the sample SD)	estimates	σ (the population SD)

p. 442

Making Connections

Identify the population, parameter, sample, and statistic in each of the following settings.

(a) A high school student was interested in finding the mean annual tuition at a 4-year U.S. college. The student randomly selected 23 U.S. colleges and found a mean annual tuition of $19,800.

Population:

Parameter:

Sample:

Statistic:

Remember **s** and **p**:
statistics come from **s**amples and
parameters come from **p**opulations

Identify the population, parameter, sample, and statistic in each of the following settings.

(a) A high school student was interested in finding the mean annual tuition at a 4-year U.S. college. The student randomly selected 23 U.S. colleges and found a mean annual tuition of $19,800.

Population: All 4-year U.S. Colleges

Parameter: μ = mean annual tuition

Sample: The 23 selected colleges

Statistic: $\bar{x}$ = the mean annual tuition in the sample = $19,800

AP® Exam Tip

Many students lose credit on the AP® Statistics exam when defining parameters because their description refers to the sample instead of the population or because the description isn't clear about which group of individuals the parameter is describing. When defining a parameter, we suggest including the word ***all*** or the word ***true*** in your description to make it clear that you aren't referring to a sample statistic.

(b) During World War II, the United States captured several tanks from the German army. Based on the serial numbers on the tanks, statisticians estimated that the German army produced 7168 tanks during the war.

Population:

Parameter:

Sample:

Statistic:

(b) During World War II, the United States captured several tanks from the German army. Based on the serial numbers on the tanks, statisticians estimated that the German army produced 7168 tanks during the war.

Population: All German tanks produced during WWII.

Parameter: The true total # of German tanks.

total

Sample: The several captured tanks.

Statistic: The estimated total number of German tanks based on the sample = 7168

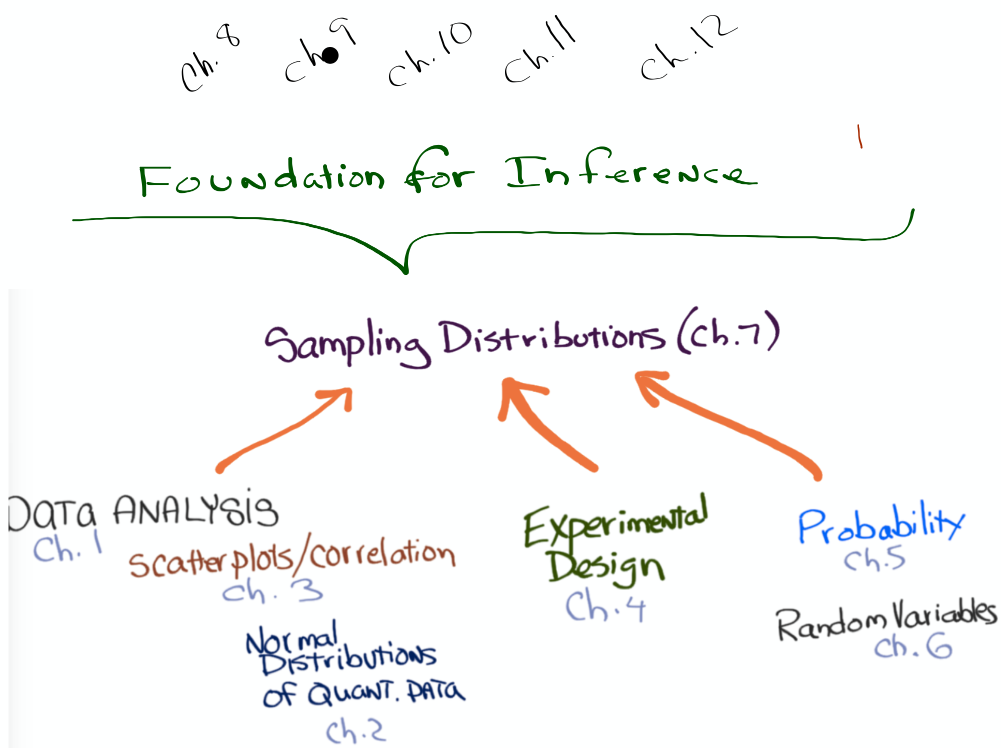
Advice from my summer conference to help students prepare for **Inference.**

> *"Don't let students get away with using μ when it should really be $\bar{x}$ or p when it should be $\hat{p}$. Being picky now will help students be successful when they get to the formulas for confidence intervals and significance tests in Ch. 8-12."*

The Big Picture: Where Chapter 7 Fits

Chapters 5 to 7 cover much of AP Statistics Topic Outline III: Anticipating Patterns: Exploring random phenomena using probability and simulations.

Chapter 5	Probability: What Are the Chances?
Chapter 6	Random Variables
Chapter 7	Sampling Distributions



ESSENTIAL QUESTIONS *How far will our estimates typically vary from the truth? What values of a statistic should be considered unusual?*

PACING **7 days**

Chapter 7: Sampling Distributions

7.1 What is a Sampling Distribution?	2 Days	T, W
7.2 Sample Proportions	1 Day	Th
7.3 Sample Means	2 Days	F M
Review, FRAPPY, and Test	2 Days	T W

Read p.440 Intro Together

Mr. Cedarlund should have been having us "buid" this over the last week, but he did not.

Instead....

We'll peek in on some data from another class.

From a huge population of pennies

A person selects a penny, looks at the date, records the age, and puts the penny back.

This is repeated over and over for days.

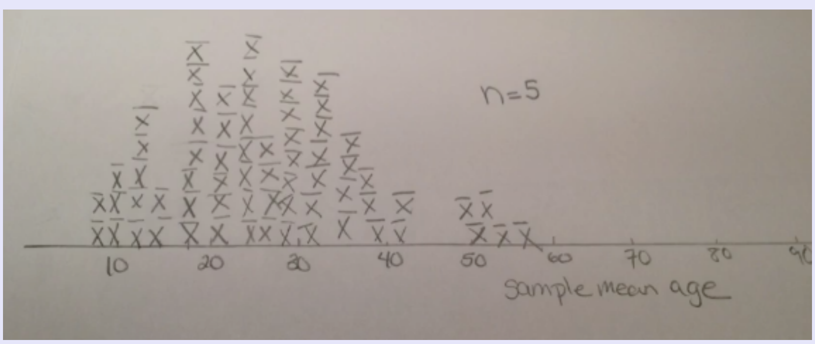
Gives us an idea of the Distribution of the Population.

Population distribution of age of penny

Next

Each person selects a SRS of 5 pennies, finds the average age, $\bar{x}$, plots $\bar{x}$.

(Returns the pennies.)



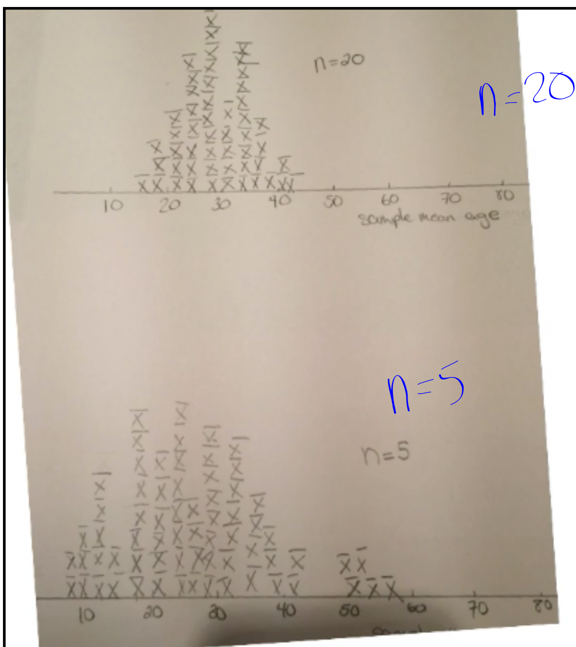
$n = 5$

Simulated Sampling Distribution of $\bar{x}$ ← of the sample means

Next each person selects an SRS of 20 pennies,

finds the sample mean, $\bar{x}$.

$n = 20$



n = 20

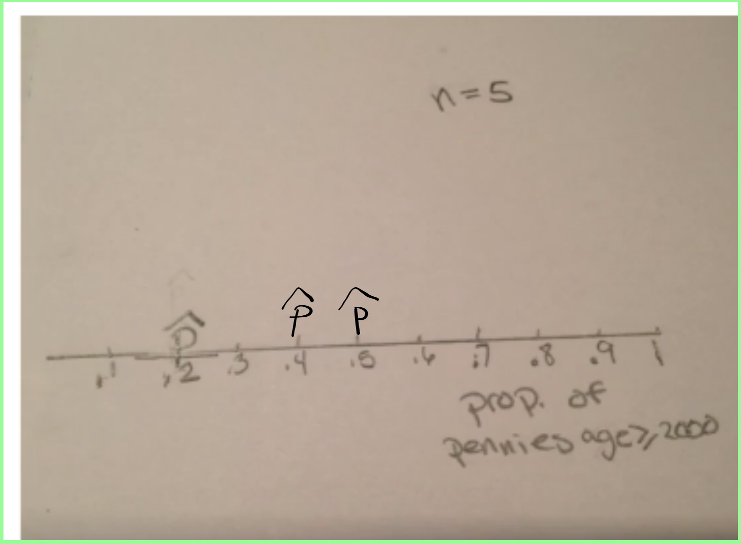
n = 5

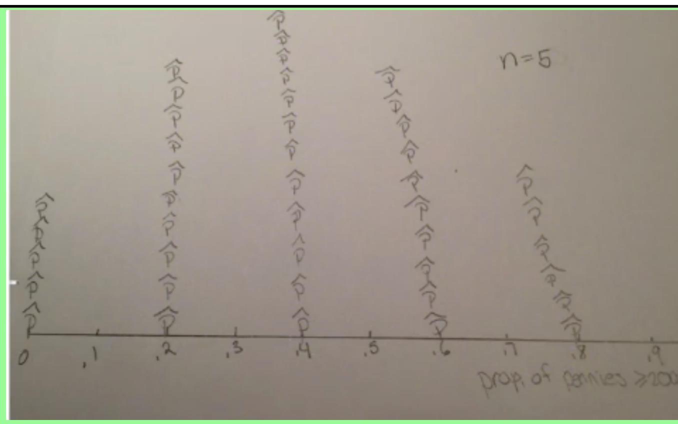
- ✓ All three distributions have the same center.
- ✓ The simulated sampling distributions have a smaller spread than the population distribution.
- ✓ As the sample size increases the variability decreases.
- ✓ The shape of the sampling distribution is still skewed right when n = 5, but less so when n = 20.

Part 2

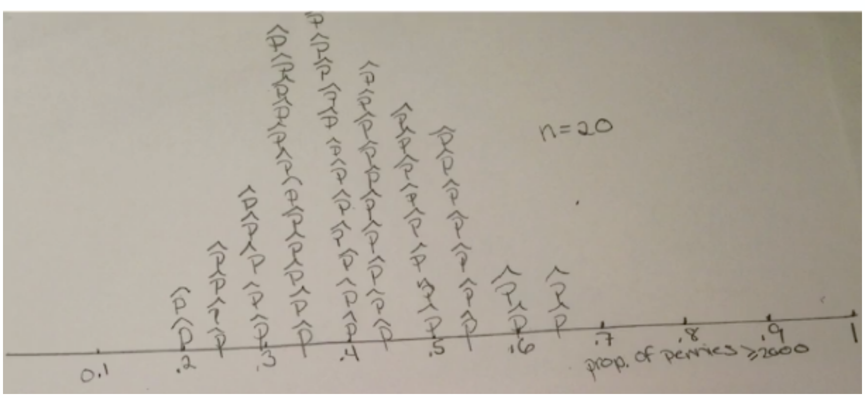
Each person takes samples of 5 pennies, records the proportion of the 5 pennies that were made in the 2000's ← called a sample proportion.

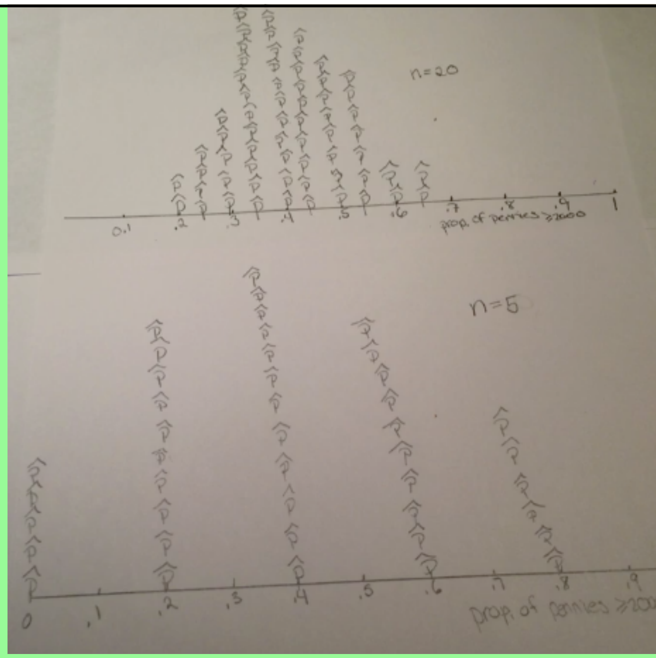
$\hat{p}$


n=5
.1 .2 .3 .4 .5 .6 .7 .8 .9 1
prop. of
pennies age≥2000


n=5
0 .1 .2 .3 .4 .5 .6 .7 .8 .9
prop. of pennies ≥200

Everyone
Repeats for $n=20$





n=20

n=5

Have We Seen Something like this before?

The two most common statistics

Sample proportions 7.2

Sample means 7.3

7.2
7.3

- Two most common statistics:
  - Sample proportion (Section 7.2)
  - Sample mean (Section 7.3)
- Goal is to have students be able to anticipate the shape, center, and variability of a sampling distribution without doing a simulation.

Sampling Distributions

Sampling variability refers to the fact that different random samples of the same size from the same population produce different values for a statistic.

Lesson 7.1: Day 1: What was the average for the Chapter 6 test?

How did the Chapter 6 test go? Today, we will be taking a **sample** from a **population**. We will use the average from the **sample** to estimate the average for the **population**.

Let's start with a very simple example. My 5th hour is very small. There were only 4 people who took the chapter 6 test. Their scores were: 60 70 80 90.

1. Make a dotplot of the population distribution.

Lesson 7.1: Day 1: What was the average for the Chapter 6 test?

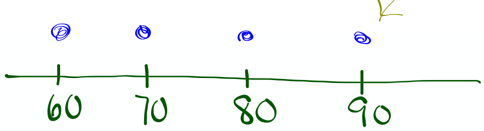
How did the Chapter 6 test go? Today, we will be taking a **sample** from a **population**. We will use the average from the **sample** to estimate the average for the **population**.

Let's start with a very simple example. My 5th hour is very small. There were only 4 people who took the chapter 6 test. Their scores were: 60 70 80 90.

1. Make a dotplot of the population distribution.

shows all individuals

1 individual with a score of 90



2. Take a sample of any 2 of the scores. Find the mean of your sample.

for example 60 and 70

$\bar{x} = \frac{60+70}{2} = 65$ ↑ sample mean

3. Figure out all of the possible samples of size 2. Calculate a sample mean for each sample of 2.

60, 70 $\bar{x} = 65$

60, 80 $\bar{x} = 70$

60, 90 $\bar{x} = 75$

70, 80 $\bar{x} = 75$

70, 90 $\bar{x} = 80$

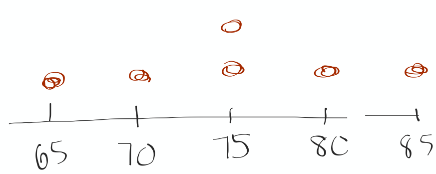
80, 90 $\bar{x} = 85$

How many combinations are there?

$_4C_2 = 6$

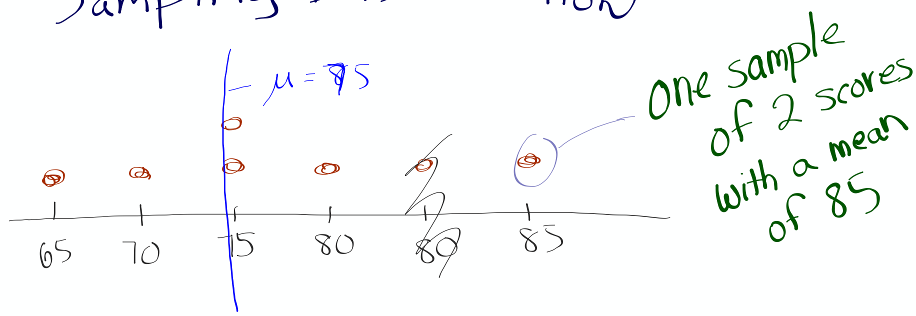
4. Make a dotplot using each of the means you found in #3.

Sampling Distribution



5. What is the mean of the population? Label this on the dotplot above.

Sampling Distribution



5. What is the mean of the population? Label this on the dotplot above.

$$\text{pop. mean} = \mu = \frac{60+70+80+90}{4} = 75$$

What is a Sampling Distribution?

Important ideas:

What is a Sampling Distribution?

Important ideas:

PARAMETERS & STATISTICS

Sampling Distribution

What is a Sampling Distribution?

Important ideas:

PARAMETERS & STATISTICS

Sampling Distribution

STATISTIC • Number that describes a sample.

Parameter • Number that describes a pop.

	means	~~pop~~ proportions	SD
Param.	μ	p	σ
Stat			

What is a Sampling Distribution?

Important ideas:

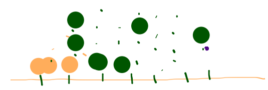
PARAMETERS & STATISTICS

STATISTIC • Number that describes a sample.

Parameter • Number that describes a pop.

	means	Propor pop	SD
Param.	μ	p	σ
Stat	$\bar{x}$	$\hat{p}$	s

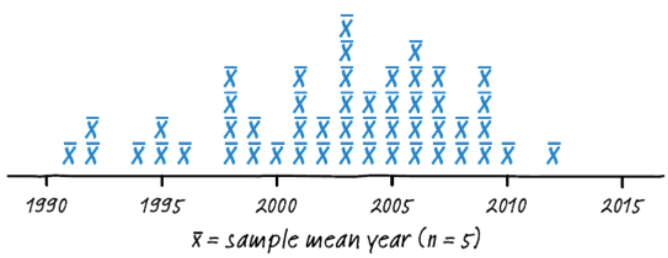
Sampling Distribution
- Shows that statistic found in all possible samples of size n



← Sampling Distrib. of statistic

Population Distrib shows all individuals

Sampling variability refers to the fact that different random samples of the same size from the same population produce different values for a statistic.



The **sampling distribution** of a statistic is the distribution of values taken by the statistic in all possible samples of the same size from the same population.

example
to look at

The Idea of a Sampling Distribution

We normally can't get the full population distribution

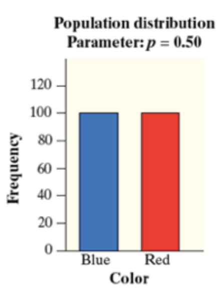


FIGURE 7.2 The idea of a sampling distribution is to take many samples from the same population, collect the value of the statistic from all the samples, and display the distribution of the statistic. The dotplot shows the approximate sampling distribution of $\hat{p}$ = the sample proportion of red chips.

The Idea of a Sampling Distribution

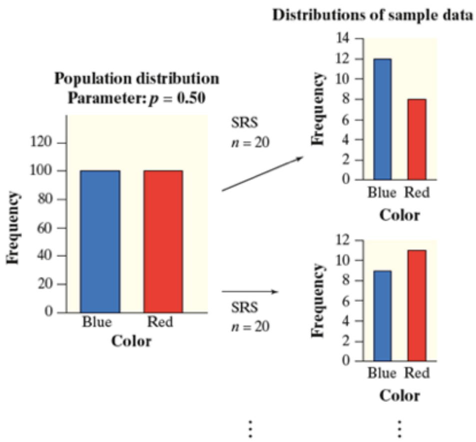


FIGURE 7.2 The idea of a sampling distribution is to take many samples from the same population, collect the value of the statistic from all the samples, and display the distribution of the statistic. The dotplot shows the approximate sampling distribution of $\hat{p}$ = the sample proportion of red chips.

The Idea of a Sampling Distribution

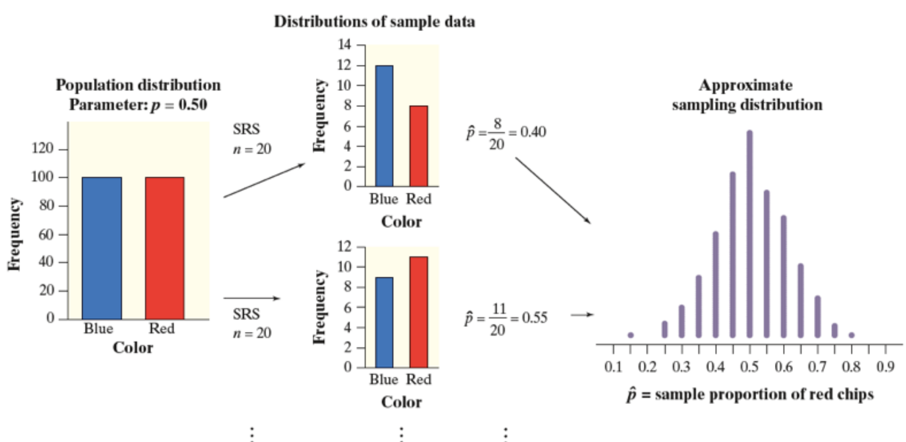
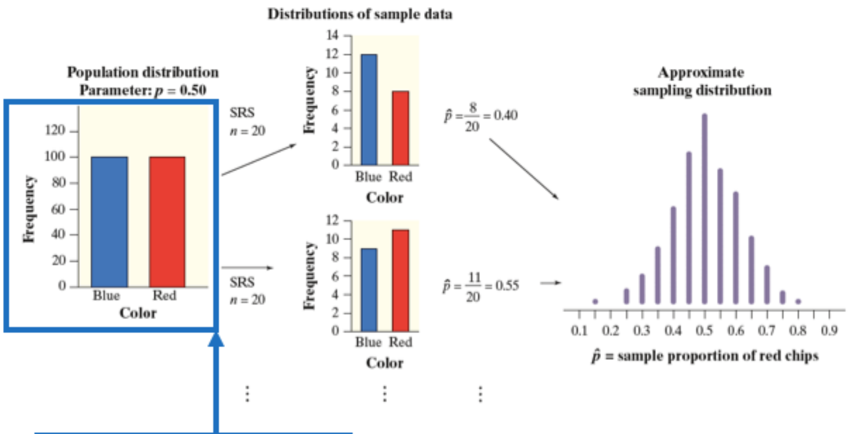
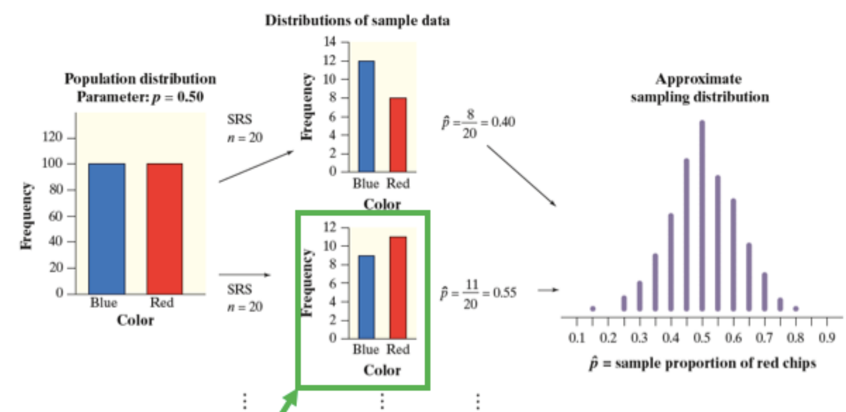


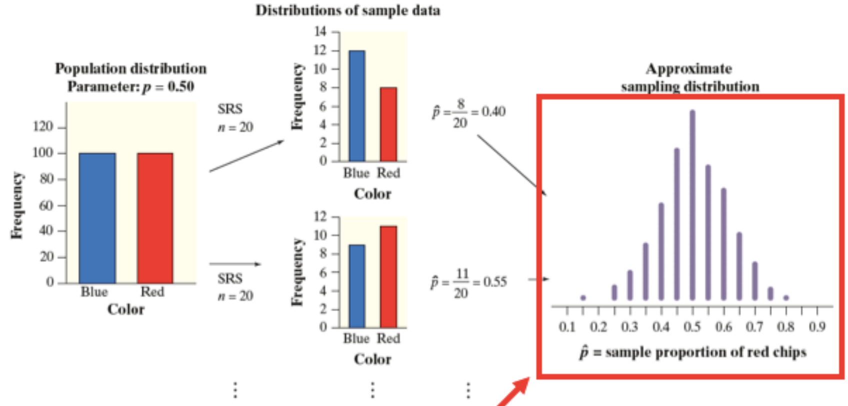
FIGURE 7.2 The idea of a sampling distribution is to take many samples from the same population, collect the value of the statistic from all the samples, and display the distribution of the statistic. The dotplot shows the approximate sampling distribution of $\hat{p}$ = the sample proportion of red chips.



- The *population distribution* gives the values of the variable for all individuals in the population.
- The *distribution of sample data* shows the values of the variable for the individuals in a sample.
- The *sampling distribution of the sample proportion* displays the values of $\hat{p}$ from all possible samples of the same size.



- The *population distribution* gives the values of the variable for all individuals in the population.
- The *distribution of sample data* shows the values of the variable for the individuals in a sample.
- The *sampling distribution of the sample proportion* displays the values of $\hat{p}$ from all possible samples of the same size.



- The *population distribution* gives the values of the variable for all individuals in the population.
- The *distribution of sample data* shows the values of the variable for the individuals in a sample.
- The *sampling distribution of the sample proportion* displays the values of $\hat{p}$ from all possible samples of the same size.

AP® Exam Tip

Terminology matters. Never just say "the distribution." Always say "the distribution of [blank]," being careful to distinguish the distribution of the population, the distribution of sample data, and the sampling distribution of a statistic.

Likewise, don't use ambiguous terms like "**sample distribution**," which could refer to the **distribution of sample data** or to the sampling distribution of a statistic. You will lose credit on free response questions for misusing statistical terms.

HOMEWORK: To determine how much homework time students will get in class, Mrs. Lin has a student select an SRS of 20 chips from a large bag. The number of red chips in the SRS determines the number of minutes in class students get to work on homework. Mrs. Lin claims that there are 200 chips in the bag and that 100 of them are red. When Jenna selected a random sample of 20 chips from the bag (without looking), she got 7 red chips. Does this provide convincing evidence that less than half of the chips in the bag are red?

1. Identify the population, parameter, sample and statistic.

Population: 200 chips Parameter: $p = \frac{100}{200} = 0.50$

Sample: 20 rand. selected chips Statistic: $\hat{p} = 7/20 = .35$

2. What is the evidence that less than half of the chips in the bag are red?

HOMEWORK: To determine how much homework time students will get in class, Mrs. Lin has a student select an SRS of 20 chips from a large bag. The number of red chips in the SRS determines the number of minutes in class students get to work on homework. Mrs. Lin claims that there are 200 chips in the bag and that 100 of them are red. When Jenna selected a random sample of 20 chips from the bag (without looking), she got 7 red chips. Does this provide convincing evidence that less than half of the chips in the bag are red?

1. Identify the population, parameter, sample and statistic.

Population: 200 chips Parameter: $p = 100/200 = 0.5$ ←

Sample: 20 randomly chosen chips Statistic: $\hat{p} = 7/20 = 0.35$

2. What is the evidence that less than half of the chips in the bag are red?

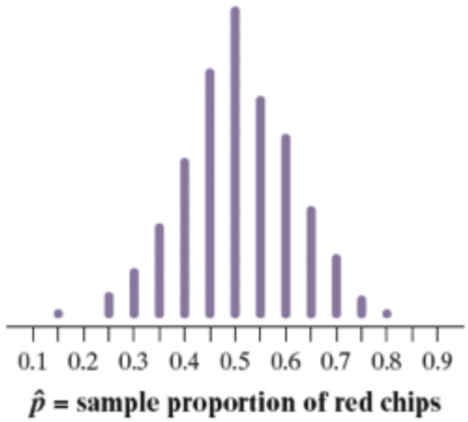
The sample had only 7 of 20 chips (35%) that were red. This is less than 50%.

3. Provide two explanations for the evidence described in part (a).

(A) We got 35% by chance

(B) There <u>are</u> less than 50% red chips in the bag.

We used technology to simulate choosing 500 SRSs of size n = 20 from a population of 200 chips, 100 red and 100 blue. The dotplot shows $\hat{p}$ = the sample proportion of red chips for each of the 500 samples.



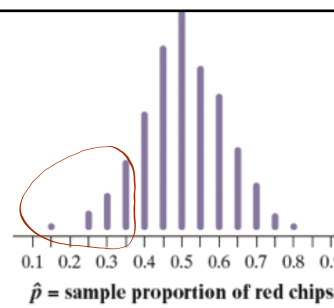
Is this...

a Population Distrib ?

a Distrib. of a sample

a Sampling Distrib.

We used technology to simulate choosing 500 SRSs of size n = 20 from a population of 200 chips, 100 red and 100 blue. The dotplot shows $\hat{p}$ = the sample proportion of red chips for each of the 500 samples.



4. There is one dot on the graph at 0.80. Explain what this value represents.

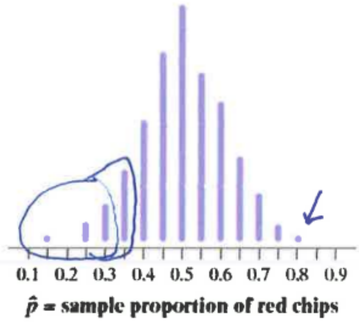
   In a SRS of 20 chips, 80% were red (16 out 20)

   (.8)(20)

5. Would it be surprising to get a sample proportion of $\hat{p}$ = 7/20 = 0.35 or smaller in an SRS of size 20 when p = 0.5? Justify your answer.

6. Based on your previous answers, is there convincing evidence that less than half of the chips in the large bag are red? Explain your reasoning.

We used technology to simulate choosing 500 SRSs of size n = 20 from a population of 200 chips, 100 red and 100 blue. The dotplot shows $\hat{p}$ = the sample proportion of red chips for each of the 500 samples.



4. There is one dot on the graph at 0.80. Explain what this value represents.

This dot represents one sample of 20 chips that contained 16 red chips.
↖ (0.8)(20)

5. Would it be surprising to get a sample proportion of $\hat{p}$ = 7/20 = 0.35 or smaller in an SRS of size 20 when p = 0.5? Justify your answer.

No. A sample of 35% or less happens relatively often.

6. Based on your previous answers, is there convincing evidence that less than half of the chips in the large bag are red? Explain your reasoning.

No. There is a large enough chance that we get a sample proportion of 0.35 or less purely by chance. → it's greater than 5%

See your Ch. 6 Test

When calculating probabilities

$$P(X=5) = \text{binompdf}(\underset{n}{17}, \underset{p}{1/6}, \underset{k}{5})$$

Need this descriptor

E -

tiny

P(Z < -1.5)

m/c .3

FR .7

7.1 ..... 1-9 (odds). Study pp. 442-447