Name_____________________________________

Chapter 7 Learning Objectives	**Section**	**Related Example on Page(s)**	**Relevant Chapter Review Exercise(s)**	**Can I do this?**
Distinguish between a parameter and a statistic.	7.1	425	R7.1	
Use the sampling distribution of a statistic to evaluate a claim about a parameter.	7.1	427	R7.5, R7.7	
Distinguish among the distribution of a population, the distribution of a sample, and the sampling distribution of a statistic.	7.1	Discussion on 428	R7.2	
Determine whether or not a statistic is an unbiased estimator of a population parameter.	7.1	Discussion on 430–431; 435	R7.3	
Describe the relationship between sample size and the variability of a statistic.	7.1	432	R7.3	
Find the mean and standard deviation of the sampling distribution of a sample proportion $\hat{p}$. Check the 10% condition before calculating $\sigma_{\hat{p}}$.	7.2	445	R7.4	
Determine if the sampling distribution of $\hat{p}$ is approximately Normal.	7.2	445	R7.4	
If appropriate, use a Normal distribution to calculate probabilities involving $\hat{p}$.	7.2	445	R7.4, R7.5	
Find the mean and standard deviation of the sampling distribution of a sample mean $\bar{x}$. Check the 10% condition before calculating $\sigma_{\bar{x}}$.	7.3	452	R7.6	
Explain how the shape of the sampling distribution of $\bar{x}$ is affected by the shape of the population distribution and the sample size.	7.3	457	R7.6, R7.7	
If appropriate, use a Normal distribution to calculate probabilities involving $\bar{x}$.	7.3	455, 459	R7.6, R7.7	

7.1 Sampling Distributions Read 424–425

There are 12 box lids. Each one has instructions, some objects, some graph paper, and a marker. Follow the instructions and record your data here.

Population: white poker chips
What 3 numbers did you select? _____ _____ _____

Population: blue and red poker chips
What 3 numbers did you select? _____ _____ _____

Population: Tub of pennies
What 3 numbers did you select? _____ _____ _____

Population: Set of 2 blue 20-sided dice
What 2 numbers did you roll? _____ _____

Population: Set of 6 blue 20-sided dice
What 6 numbers did you roll? _____ _____ _____ _____ _____ _____

Population: Set of 4 red 6-sided dice
What 4 numbers did you roll? _____ _____ _____ _____

Population: Set of 4 green 6-sided dice
What 4 numbers did you roll? _____ _____ _____ _____

Population: Deck of Red Cards
How many face cards did you get among your 5 cards? _____

Population: Deck of Blue Cards
How many hearts did you get among your 5 cards? _____

Population: Set of 5 red 4-sided dice
How many 2s did you roll? _____

Population: Set of 5 blue 4-sided dice
How many 2s did you roll? _____

Population: Set of 5 white 6-sided dice
How many 2s did you roll? _____

***

Recall: How do you identify Categorical vs. Quantitative?

What is a parameter? What is a statistic?

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

How is one related to the other?

Identify the population, the parameter, the sample, and the statistic:
(a) A guidance counselor wants to know the 25th percentile for the distribution of gpa of high school students, so she takes a sample of 60 students and calculates Q1 = 2.41.

(b) A Pew Research Center Poll asked 1009 13- to 17-year-olds in the United States if they have a smart phone. Of the respondents, 73% said “Yes.”

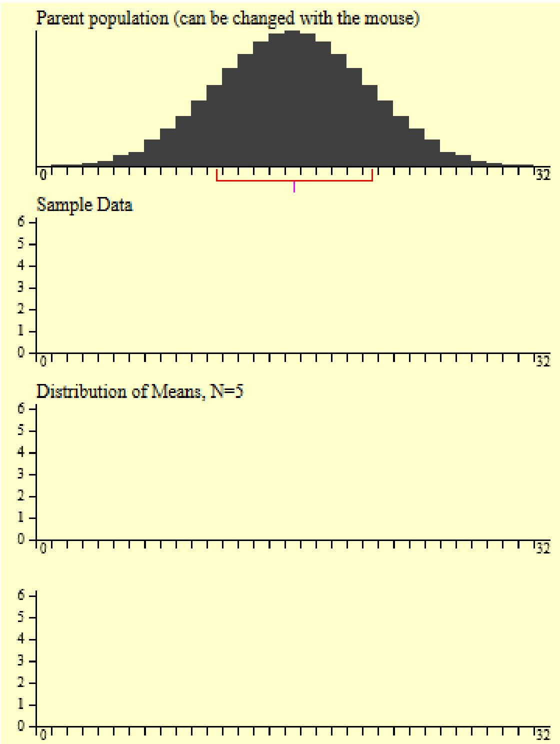
Read 425–429
What is sampling variability?

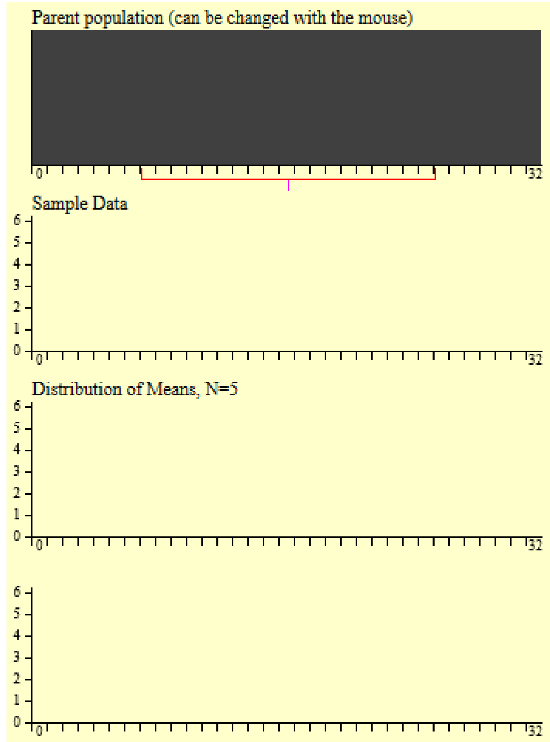
sampling variability: the value of a statistic varies in repeated sampling

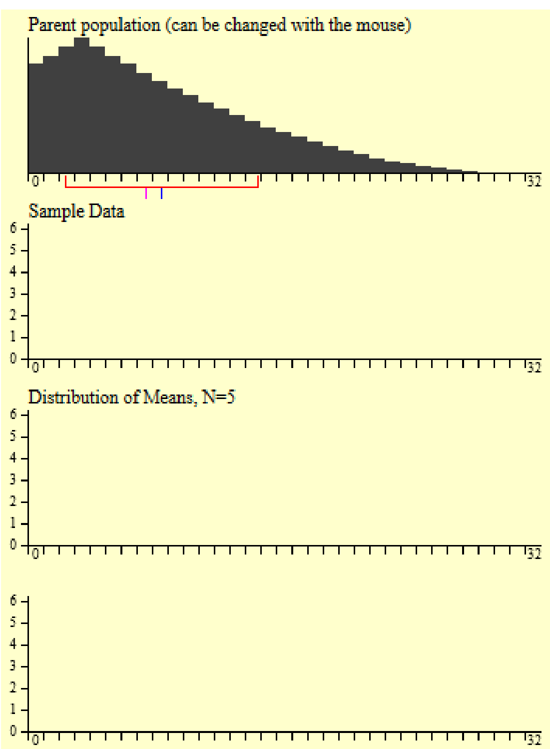
What is a sampling distribution?

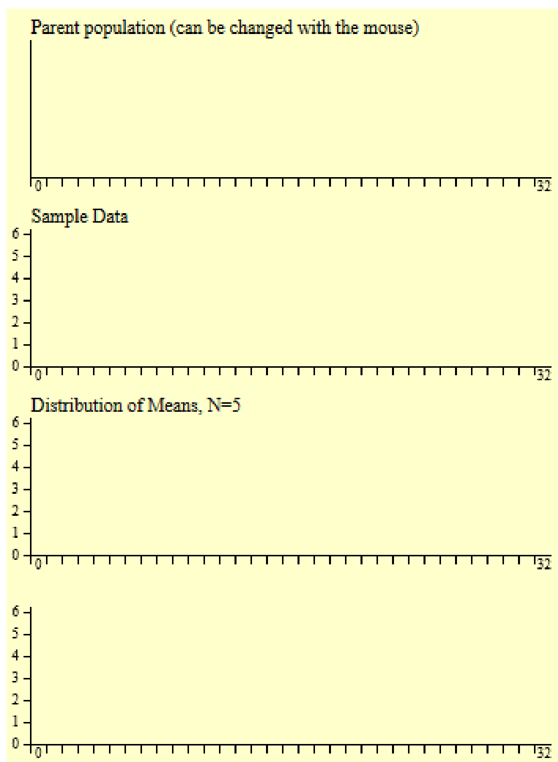
The **sampling distribution** of a statistic is the distribution of values taken by the statistic in all possible samples of the same size from the same population.

What is the difference between the distribution of the population, the distribution of the sample, and the sampling distribution of a sample statistic?







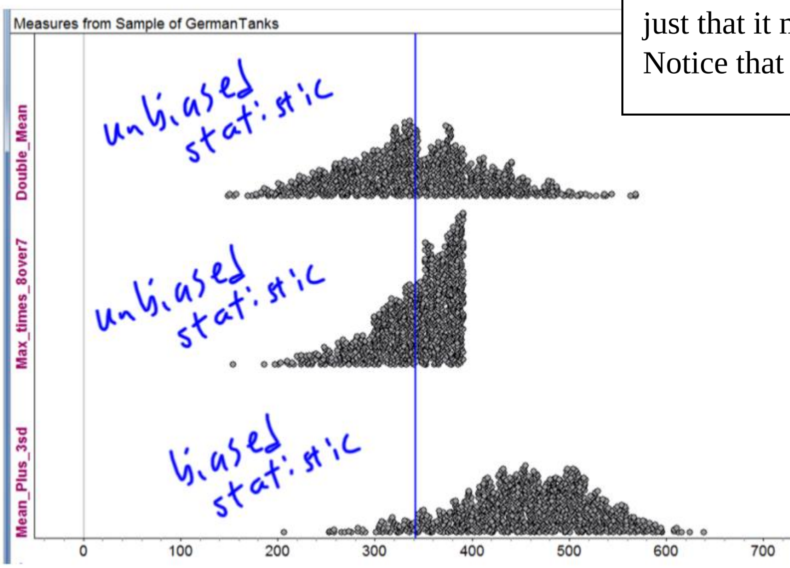


Read 429–435
What is an unbiased estimator? What is a biased estimator?

A statistic used to estimate a parameter is an **unbiased estimator** if the mean of its sampling distribution is equal to the value of the parameter being estimated.

Examples:

We don't say a statistic is always biased,
just that it might be a biased estimator OF some parameter.
Notice that this is NOT about biased sampling.



How can you reduce the variability of a statistic?
What effect does the size of the population have on the variability of a statistic?

The **variability of a statistic** is described by the spread of its sampling distribution. This spread is determined mainly by the size of the random sample. Larger samples give smaller spreads. The spread of the sampling distribution does not depend much on the size of the population, as long as the population is at least 10 times larger than the sample.

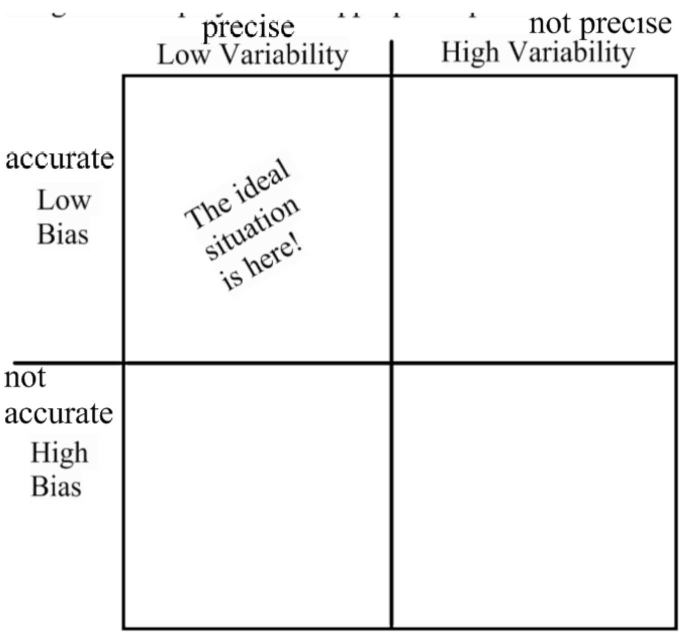
So this is our "10% rule": $n \leq \frac{1}{10}N$ or, in words, our sample is no more than 10% of the population, when sampling without replacement. This rule isn't needed when we sample with replacement, nor in experiment situations. We can either say that our sample is no more than 10% of the population, or we can say that the population is at least 10 times larger than our sample.

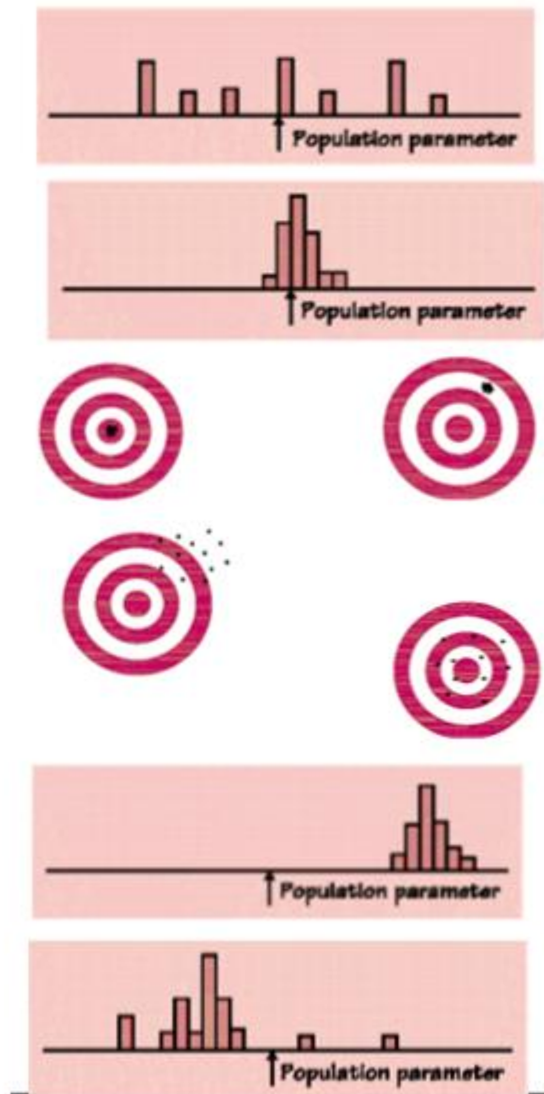
Suppose I ask the each person in class to estimate the height of the light posts in the parking lot. What is the difference between accuracy and precision? How does this relate to bias and variability?

If everyone's estimate is close to the actual height, then the estimates have high ________ or low ________

If everyone's estimate is close to one another's, then the estimates have high ________ or low ________

Draw arrows or lines from the images at the right to the appropriate boxes or label the boxes A, B, C, D and label the images for which box they go in. You might also consider just labeling the images as high bias, low bias, high variability, and low variability.





HW #1: page 436 (1–19 odd)
7.2 Sampling Distribution of a Sample Proportion
Read 440–443

In the context of the Candy Machine Applet, explain the difference between the distribution of the population, the distribution of a sample, and the sampling distribution of the sample proportion.

Based on the Candy Machine Applet and the Penny Activity, describe what we know about the shape, center, and spread of the sampling distribution of a sample proportion.

When is it OK to say that the distribution of $\hat{p}$ is approximately Normal?

Read 444–445

What are the mean and the standard deviation of the sampling distribution of a sample proportion? Are these formulas on the formula sheet? Are there conditions that need to be met for these formulas to work?

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: $n \leq \frac{1}{10}N$.
- 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 $n(1-p) \geq 10$.

Read 445–446

In a large corporation, 82% of all employees are planning to sign up for the "Plan A" insurance plan. What is the probability that an SRS of size 125 will give a sample proportion of at most 78%?

HW #2: page 436 (21–24), page 447 (29-39 odd)
7.3 Sampling Distribution of a Sample Mean
Based on the penny activity, what do we know about the shape, center, and spread of the sampling distribution of a sample mean?

Read 451–453

What are the mean and standard deviation of the sampling distribution of a sample mean? Are these formulas on the formula sheet? Are there any conditions for using these formulas?

Suppose that $\bar{x}$ is the mean of an SRS of size n drawn from a large population with mean μ and standard deviation σ. Then:

- The **mean** of the sampling distribution of $\bar{x}$ is $\mu_{\bar{x}} = \mu$.
- The **standard deviation** of the sampling distribution of $\bar{x}$ is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

as long as the *10% condition* is satisfied: $n \leq \frac{1}{10}N$.

Read 453–456

What is the shape of the sampling distribution of a sample mean when the sample is taken from a Normally distributed population? Does the sample size matter?

SAMPLING DISTRIBUTION OF A SAMPLE MEAN FROM A NORMAL POPULATION

Suppose that a population is Normally distributed with mean μ and standard deviation σ. Then the sampling distribution of $\bar{x}$ has the Normal distribution with mean μ and standard deviation (provided the 10% condition is met) $\sigma/\sqrt{n}$.

At a corn chip manufacturer, chips are placed in bags by a machine.
The distribution of weights in the bags is approximately Normal, with a mean of 24.1 ounces and a standard deviation of 0.16 ounces.
(a) Without doing any calculations, explain which outcome is more likely: randomly selecting a single bag and finding that the contents weigh less than 24 ounces or randomly selecting 10 bags and finding that the average contents weigh less than 24 ounces.

(b) Find the probability of each event described above.

Read 457–460
What is the shape of the sampling distribution of a sample mean when the sample is NOT taken from a Normally distributed population? Does the sample size matter? Does this concept have a name?

DEFINITION: Central limit theorem (CLT)

Draw an SRS of size n from any population with mean μ and finite standard deviation σ. The **central limit theorem (CLT)** says that when n is large, the sampling distribution of the sample mean $\bar{x}$ is approximately Normal.

NORMAL/LARGE SAMPLE CONDITION FOR SAMPLE MEANS

- If the population distribution is Normal, then so is the sampling distribution of $\bar{x}$. This is true no matter what the sample size n is.
- If the population distribution is not Normal, the central limit theorem tells us that the sampling distribution of $\bar{x}$ will be approximately Normal in most cases if $n \geq 30$.

Suppose that the mean household income in a certain state follows a right-skewed distribution with a mean of \$53000 and a standard deviation of \$9000. How likely is it that a random sample of 100 households in that state will have a total of income of at least \$5,100,000?

Don't need to do the four-step process on this one!

	Categorical	Quantitative
Center	$\mu_{\hat{p}} = p$, <u>if</u> a SRS.	$\mu_{\bar{x}} = \mu$, <u>if</u> a SRS.
Spread	If n≤10% of N, then $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$	If n≤10% of N, then $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
Shape	If np≥10 and n(1-p)≥10, then $\hat{p}$ has approx. a normal dist.	If n≥30, then $\bar{x}$ has approx. a normal dist. If the pop. is normal or approx. normal, then the sampling dist. of $\bar{x}$ is too (for any n).

HW #3 page 447 (41, 43–46), page 461 (49–63 odd, 65–68)
Chapter 7 Review/FRAPPY
FRAPPY: 2009 #2 (stopping distances)
HW #4 page 466 Chapter Review Exercises
Chapter 7 Review
HW #5 page 468 Chapter 7 AP Statistics Practice Test
Chapter 7 Test