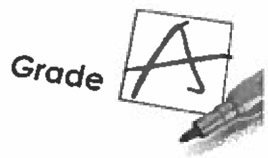
Name: ______________________________ Hour: _____ Date: _____

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.

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

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

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

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

Name: ______________________ Hour: _____ Date: ______

Lesson 7.1 Day 1– What is a Sampling Distribution?

Important ideas:

Check Your Understanding

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.

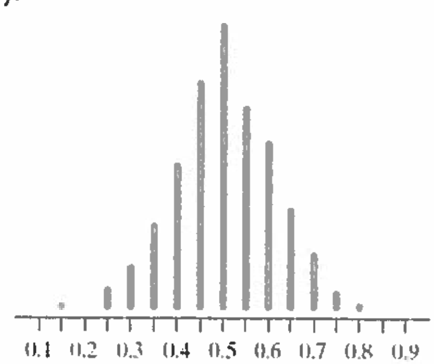
   Population: ______________________ Parameter: ______________________

   Sample: ______________________ Statistic: ______________________

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

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

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.



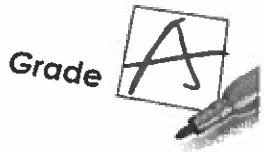
4. There is one dot on the graph at 0.80. Explain what this value represents.

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.

Name: ______________________________ Hour: _____ Date: ______

Lesson 7.1: Day 2: What was the real 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**.

Yesterday we looked at a very small class of students as the population. In reality there were many students who took the test. Take a random sample of 5 students and record their scores. Then find the mean. Repeat this for a total of 4 times.

Scores:______________ Mean:______ Scores:______________ Mean:______

Scores:______________ Mean:______ Scores:______________ Mean:______

1. Write each mean on a different sticker and put the stickers in the appropriate location on the poster at the front of the room. Copy down the dotplot that is created on the poster.

2. What does each dot on the poster represent?

3. What do you think the true Chapter 6 test average is?

4. A **sampling distribution** shows the means calculated from all of the possible samples of size 5 from the population. Is the above dotplot a sampling distribution? Explain.

5. We took a random sample of 5 midterm scores at Rockford high school and got a mean of 68. Is this convincing evidence that Rockford students did worse than students at our school?

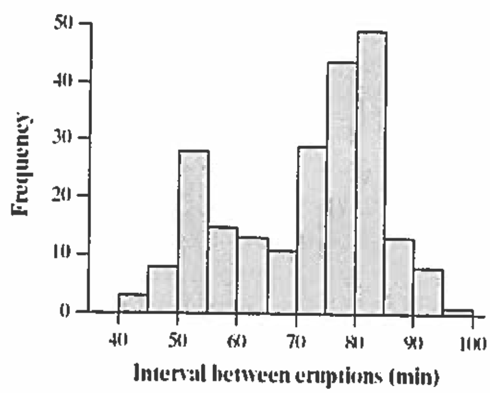
Name: ______________________________________ Hour: _____ Date: ______

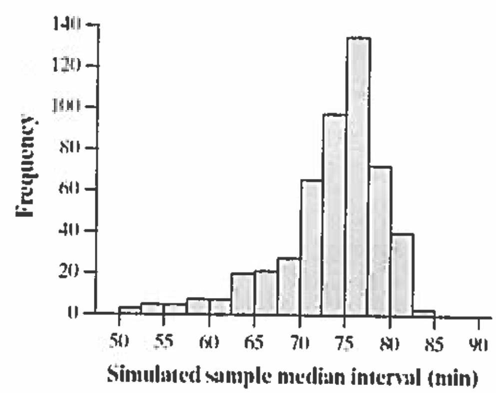
Lesson 7.1 Day 2– Biased and Unbiased Estimators

Important ideas:

Check Your Understanding

The histogram on the left shows the interval (in minutes) between eruptions of the Old Faithful geyser for all 222 recorded eruptions during a particular month. For this population, the median is 75 minutes. We used technology to take 500 SRSs of size 10 from the population. The 500 values of the sample median are displayed in the histogram on the right. The mean of these 500 values is 73.5.





1. Is the sample median an unbiased estimator of the population median? Justify your answer.

2. Suppose we had taken samples of size 20 instead of size 10. Would the variability of the sampling distribution of the sample median be larger, smaller, or about the same? Justify your answer.

3. Describe the shape of the sampling distribution of the sample median.

Name: ______________________________________ Hour: _____ Date: ______

Lesson 7.2: What's the proportion of orange Reese's Pieces?

If we take a sample of Reese's Pieces, what proportion of the candies will be orange?

Suppose a large bag of Reese's Pieces has 1000 pieces. The manufacturer says that exactly 40% of the candies are orange. If we select a sample of 50 pieces, how many will be orange? Let X = the number of orange candies in the sample.

1. What type of probability distribution does X have? Justify.

2. Draw a sample of 50 Reese's Pieces using the applet. How many pieces were orange? Repeat this 5 times. Write the values below.

3. Write the values on sticker dots and add it to the dotplot on the board. Sketch the dotplot below.

4. What does each dot represent?

5. What is the mean and the standard deviation for the distribution of X? Show work.

6. What is the approximate shape of the sampling distribution for X? Explain and sketch it below.

Name: ______________________________ Hour: _____ Date: _____

Instead of finding the number of candies that are orange, we will now find the **proportion** of candies that are orange.

7. Use your samples from #2 and turn each number of orange candies into the **proportion of orange candies** in the sample ($\hat{p}$). Write the proportions below and add them to the second dotplot on the board.

8. Sketch the dotplot below.

9. What does each dot represent?

10. Find the new mean and standard deviation. Show work.

11. What is the approximate shape of the sampling distribution for $\hat{p}$? Explain and sketch it below.

12. We know that bags of Reese's Pieces contain exactly 40% that are orange. If we select a random sample of 50 candies, what is the probability that the sample proportion will be 50% or greater?

Name: ______________________________ Hour: _____ Date: _____

Lesson 7.2 – The Sampling Distribution of $\hat{p}$

Important ideas:

Check Your Understanding

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

1. Identify the mean of the sampling distribution of $\hat{p}$.

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

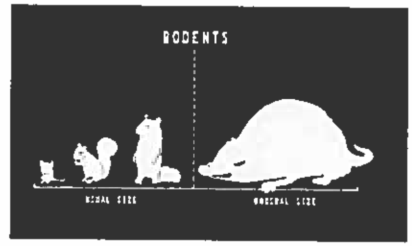
3. Is the sampling distribution of $\hat{p}$ approximately Normal? Check that the Large Counts condition is met.

4. Find the probability that the random sample of 1000 young adults will give a result within 2 percentage points of the true value.

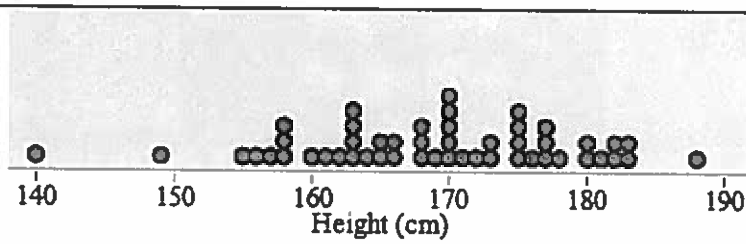
5. If the sample size were 9000 rather than 1000, how would this change the sampling distribution of $\hat{p}$?

Name: ________________________________ Hour: _____ Date: ______

Lesson 7.3: Day 1: How tall are we?



How tall are high school seniors in Michigan? Attached are the heights of all 50 high school seniors at a small high school in the upper peninsula.

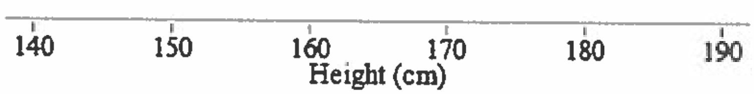


1. Make a guess at the mean of all 50 students. Make another guess of the standard deviation of all 50 students.

2. Select a random sample of 5 students and calculate the mean height for the sample. Repeat for 4 samples total.

   Heights:______________________________ $\bar{x}$ =____________
   Heights:______________________________ $\bar{x}$ =____________
   Heights:______________________________ $\bar{x}$ =____________
   Heights:______________________________ $\bar{x}$ =____________

3. Add your sample means to the dotplot on the board. Sketch it below.



4. Describe the shape, center, and variability of this dotplot.

5. Compare the two dotplots above. How are the dotplots similar? How are they different?

Name: ______________________________ Hour: _____ Date: _____

Lesson 7.3 Day 1 – Sample Means

Important ideas:

Check Your Understanding

The length of human pregnancies from conception to birth varies according to a distribution that is approximately Normal with mean 266 days and standard deviation 16 days.

1. Find the probability that a randomly chosen pregnant woman has a pregnancy that lasts for more than 270 days.

Suppose we choose an SRS of 6 pregnant women. Let $\bar{x}$ = the mean pregnancy length for the sample.

2. What is the mean of the sampling distribution of $\bar{x}$?

3. Calculate and interpret the standard deviation of the sampling distribution of $\bar{x}$. Verify that the 10% condition is met.

4. Find the probability that the mean pregnancy length for the women in the sample exceeds 270 days.

Student	Height	Student	Height	Student	Height	Student	Height	Student	Height
1	180	**11**	162	**21**	168	**31**	161	**41**	182
2	158	**12**	170	**22**	158	**32**	180	**42**	165
3	173	**13**	149	**23**	165	**33**	156	**43**	163
4	175	**14**	181	**24**	140	**34**	183	**44**	160
5	170	**15**	157	**25**	166	**35**	172	**45**	163
6	168	**16**	166	**26**	158	**36**	175	**46**	173
7	176	**17**	163	**27**	168	**37**	169	**47**	170
8	175	**18**	170	**28**	171	**38**	182	**48**	164
9	177	**19**	177	**29**	188	**39**	177	**49**	175
10	178	**20**	155	**30**	163	**40**	183	**50**	170

Name: ______________________________ Hour: _____ Date: ______

Lesson 7.3: Day 2: Who has better ACT scores?

VS

The ACT test is scored with whole numbers from 0 to 36. We will use the website www.tinyurl.com/EKstats66 to take samples of ACT scores from EK and Rockford.

Click "Begin" and you will see the population distribution of ACT scores from EK.

1. Describe the shape, center, and variability of the distribution of ACT scores for EK.

2. Click "Animated" to take a sample of 5 ACT scores.

List 5 estimated scores here:________________ Estimated Mean (blue box):________

Click "Animated" several more times. Then click "10,000" to take 10,000 samples of size 5.

3. The blue boxes make the sampling distribution of $\bar{x}$. How do we know that the sampling distribution of $\bar{x}$ is approximately normal (Lesson 7.3 Day 1)?

4. Now let's look at the distribution of ACT scores for Rockford. Click "Clear lower 3" and then change the distribution from "Normal" to "Skewed". What is the shape of this distribution? Why does this distribution make sense for Rockford?

Change both of the bottom two dropdown menus to "Mean". The first one should be "N=2" and the second one should be "N=25". The click "10,000" to take 10,000 samples.

5. Describe the shape of the sampling distribution of $\bar{x}$ when N = 2.

6. Describe the shape of the sampling distribution of $\bar{x}$ when N = 25.

Name: ______________________________ Hour: _____ Date: _____

Lesson 7.3 Day 2– The Central Limit Theorem

Important ideas:

Check Your Understanding

Keith is the manager of an auto-care center. Based on service records of 3500 customers from the past year, the time (in hours) that a technician requires to complete a standard oil change and inspection follows a right-skewed distribution with $\mu = 30$ minutes and $\sigma = 20$ minutes. For a promotion, Keith randomly selects 40 current customers and offers them a free oil change and inspection if they redeem the offer during the next month. Keith budgets an average of 35 minutes per customer for a technician to complete the work. Will this be enough?

(a) Describe the shape of the sampling distribution of $\bar{x}$ for samples of 40 randomly selected customers. Justify your answer.

(b) Find the mean and standard deviation of the sampling distribution of $\bar{x}$. Be sure to check the 10% condition.

(c) Calculate the probability that the average time it takes to complete the work exceeds 35 minutes.

(d) How much average time per customer should Keith budget if he wants to be 99% certain that he doesn't go "over budget"?

Name: ______________________________ Hour: _____ Date: _____

Chapter 7 Review

A number that describes the whole population is known as a ______________.

A number that is calculated from a sample is known as a ______________.

We always use a ____________ to estimate a ______________.

In Section 7-2, we used a _________________ to estimate a population proportion.

In Section 7-3, we used a _________________ to estimate a population mean.

Summary:

	Sample Proportions	Sample Means
What is the parameter?		
What is the statistic?		
Draw Sampling Distribution.		
When is the sampling distribution approximately normal?		
What is the mean of the sampling distribution?		
What is the standard deviation of the sampling distribution?		
What condition must be satisfied in order to use the above formula?		
What is the formula for a z-score?		

Old stuff from Chapter 6: Binomial Distributions

The German tank problem

Imagine that you want to know how many taxis there are in London or New York. Every licensed taxi carries a number. You obviously will not see all the taxis in town, but in the course of five minutes you might see four. Suppose that one bears the number 17, one is 250, one is 337, one is 591. Can you estimate how many taxis there are in all?

It will have to be more than 591. (We assume the numbers start at 1.) Might it be as high as 1000? If so, we might have expected one of our numbers to be in the 700s or higher. It probably will not be as high as 10 000: if it were, we would have expected to see some four-digit numbers in our sample.

In fact a statistical formula gives the most likely answer. It is shown in the box.

The only time we really want to know how many London taxis there are is when it is raining. There was a time, though, when a similar problem was of desperate importance.

In 1944 the Allies were preparing for D-Day. They needed to know how many German tanks they might face on the Normandy beaches. Specifically, they were worried about the new Mark V "Panther" tank; it was heavier, with a bigger, longer gun, and it was feared that it could outperform the American Shermans. It had recently been encountered in Italy. The hope was that, since it was new, there would only be a few of them in France.

Hope was not enough. Planners had to know exactly how many Panthers there were.

Tanks, like taxis, carry numbers – dozens of them. They have chassis numbers, engine numbers, gun-barrel numbers. And some of those numbers run in sequence, from 1 to however many tanks have been made.

Tanks, though, are harder to catch than taxis. The allies had captured just two Mark Vs: one in Sicily, one in Russia. Would two be enough to do the taxi trick and work out how many tanks there were in all? The problem was handed to American statisticians.

The chassis numbers did not help. They knew, from other tanks, that chassis were made by five different manufacturers with big breaks in the number sequences. Gearboxes, on the other hand, were numbered in an unbroken sequence from 1 on up. Even so, two is a pretty small sample. Still, the formula gave an answer.

Fortunately, tanks have bogie wheels that support the tracks; the bogie wheels have rubber tyres; the tyres are made on a mould; and each mould bears a number that also gets moulded onto the tyre. Nor do moulds last for ever; the analysts asked British tyre manufacturers how many tyres they would expect a mould to make before it is replaced with a new-numbered one. More fortunately still, each Panther tank had eight axles, and each axle had six bogie wheels, making 48 wheels per tank.

They applied the taxi formula, suitably adjusted, to their 96 differently numbered bogie tyres. They came up with an answer; and it agreed very well with their gearbox answer. The Germans, they estimated, were producing Panther tanks at the rate of 270 a month. This was many more than the D-day planners had expected. The assault plan was revised. The landings succeeded.

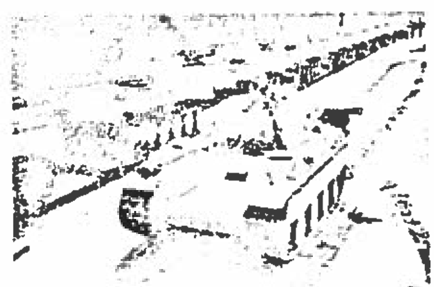
And were the estimates accurate? After the war the exact figures were found. In February 1944 276 Panther tanks had been produced. The estimate had been 270. The statisticians had got it almost exactly right.

Julian Champkin

The formula to estimate the total number of taxis (or tanks) is

$N = m - 1 + m/k$

where m is the highest serial number that we have spotted and k is the number of taxis or tanks we have seen. In our example, of four taxis with 591 as the highest serial number, this gives an answer of 738 taxis in London. Spotting more than four taxis would improve our estimate; and more complex formulae, giving Bayesian probability distributions, are available.



Panzer III "Panther". German Federal Archive

GERMAN TANK PROBLEM Names________________

During World War II, Allied intelligence reports on Germany's production of tanks varied widely and were somewhat contradictory. Statisticians set to work on improving the estimates. In 1943, they developed a method that used the information contained in the serial numbers stamped on the tanks. They discovered that the serial numbers on the tanks were consecutive. That is, the tanks were numbered in a manner equivalent to 1,2,3,……N. Capturing a tank was like randomly drawing an integer from this sequence.

One sample collected in 1943 found tanks with the following serial numbers:

21 123 201 297 342

What is your best estimate for the number of tanks, N? ______. Is this value a parameter or a statistic?

Describe the rule that you used to find your estimate:

List five other rules that could be used to estimate the number of tanks N.

1.

2.

3.

4.

5.

What makes a statistic a good one?

How could we figure out which rule gives the best estimate?

WHICH RULE GIVES THE BEST ESTIMATE?

List the three rules selected by the class as the best estimators:

1.

2.

3.

Each group will now be given a bag with integers from 1 to N contained in the bag. Randomly draw five numbers out of the bag and record the values. Return the numbers, mix well and repeat.

My group will be testing rule # ______. Round your estimate to the nearest whole tank.

Sample #1 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #2 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #3 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #4 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #5 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #6 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #7 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #8 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #9 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #10 ______ , ______, ______ , ______, ______ . Estimate: ______

Make a dotplot of your estimates below. Interval widths are already set up.

```
5
4
3
2
1
0 <----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o >
       18   20   22   24   26   28   30   32   34   36   38   40   42   44   46   48   50   52   54
```

What is the mean of your estimates? ________ What is the standard deviation? ________

Now combine your results along with the other groups who are testing the same rule and make a combined dotplot of estimates. Ask Mr. Wilcox for the dotplot paper.

PREDICTED NUMBER OF TANKS USING THE FOLLOWING RULE:

____

15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0 <----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o >

18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54

Mean: __________ Standard Deviation: __________

CENTRAL LIMIT THEOREM PENNIES LAB Name____________________

In this lab we will be interested in the age of the pennies (in years) and not the year the pennies were produced. To find the age of a penny, simply subtract the year on the penny from the current year.

1. Take a random sample of three pennies. Calculate the average age of the sample. This is called a sample mean. Record your result.

Sample mean: ______

2. Replace the three pennies. Then take another sample of three pennies. Find the sample mean and record it.

Sample mean: ______

3. Do this a total of five times. Record your three other sample means below:

Sample mean: ______ Sample mean: ______ Sample mean: ______

4. Write each of your sample means on a different post-it note and then take your post-it note to the histogram on the whiteboard. Record the results of the "dotplot" on the whiteboard below:

```
12
11
10
9
8
7
6
5
4
3
2
1
0  <----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o >
        0    2    4    6    8   10   12   14   16   18   20   22   24   26   28   30   32   34   36   38
```

TheStatsMedic

5. We will now change the size of the sample. Take a random sample of 10 pennies. Calculate the average age of the pennies. Record it here: ________

6. Do this a total of five times. Record your four other sample means below:

________ ________ ________ ________

7. Write each of your sample means on a different post-it note and then take your post-it note to the histogram on the whiteboard. Record the results of the histogram on the whiteboard below:

20
18
16
14
12
10
8
6
4
2
0 <----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o----o >
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38

Summary:

- If the population distribution is normal, the sampling distribution of $\bar{x}$ is also ________________. This is true for any sample size n.
- If the population distribution is NOT normal, the **Central Limit Theorem (CLT)** tells us that the sampling distribution of $\bar{x}$ will be approximately normal in most cases if ____________.

AP Statistics Activity Wrap-up

Name_______________

Activity Name:

Describe the activity or context:	
What important statistical concepts did we learn?	

Activity Name:

Describe the activity or context:	
What important statistical concepts did we learn?	

Activity Name:

Describe the activity or context:	
What important statistical concepts did we learn?	

Activity Name:

Describe the activity or context:	
What important statistical concepts did we learn?	

Activity Name:

Describe the activity or context:	
What important statistical concepts did we learn?	

Activity Name:

Describe the activity or context:	
What important statistical concepts did we learn?	