AP Stats Ch 18 - Required Practice **Name:** SOLUTIONS **Per:** ____

SAMPLING APPLET ACTIVITY – DAY 1

At least one person in your group use your phone's web browser to display: www.mrfelling.com/sa1

....

(the full activity is detailed in the filled-in notes if you were absent and need to refer to it)

....

Press the 'sample' button a few more times to see what happens as you continue to take samples of size n=9. ....

Compare the mean and standard deviation of the Sampling Distribution of Sample Means to the mean and standard deviation of the population. Notice that the means are about the same, but the standard deviation of the Sampling Distribution of Sample Means is much smaller than the population. Write a sentence or two explaining why you believe this is true:

For a sample of 9 to have a really high (or low) mean, all of the samples would need to be high (or low). This is unlikely, so the means are closer to the 'middle'.

Let's investigate......

Try this for yourself, and fill in the standard deviations for the population and for the Sampling Distribution for each experiment (remember to press the 'Sample400x' button before recording the standard deviations.)

	Population	Sampling Distribution	
Experiment 1 (n=1):	$\sigma =$ 20,	$\sigma_{\bar{X}} =$ 18.1	$\frac{20}{18.1} = 1.1 \approx 1 = \sqrt{1}$
Experiment 2 (n=4):	$\sigma =$ 20,	$\sigma_{\bar{X}} =$ 10.3	$\frac{20}{10.3} = 1.9 \approx 2 = \sqrt{4}$
Experiment 3 (n=9):	$\sigma =$ 20,	$\sigma_{\bar{X}} =$ 6.4	$\frac{20}{6.4} = 3.1 \approx 3 = \sqrt{9}$
Experiment 4 (n=16):	$\sigma =$ 20,	$\sigma_{\bar{X}} =$ 4.7	$\frac{20}{4.7} = 4.3 \approx 4 = \sqrt{16}$

Can you find any (approximate) relationship between the population and sampling distribution standard deviations? Write down what ~~you think is occurring~~:

$\frac{\sigma}{\sigma_{\bar{x}}} \approx \sqrt{n}$ or $\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$ and $\boxed{\mu_{\bar{x}} = \mu}$

Try this with the other shapes as well. How large does the sample size need to be before we can say the sampling distribution is approximately normal?

For bi-modal, shape is approx normal with n=16 or higher

For skewed, shape is approx normal with n=9 or higher

For normal, shape is approx normal with n=1 or higher

The closer the population shape is to normal, the smaller the sample size can be, and the sampling distribution is approx. normal. For very different shapes, is normal with n>25

SAMPLING APPLET ACTIVITY – DAY 2

Yesterday, we investigated the Sampling Distribution of Sample Means – which would apply whenever we have a numerical variable for which we could compute a mean. But what if we have categorical data, for example, a 'yes'/'no' situation for which we can only compute the percentage, or, *proportion,* of 'yes' in a population or sample?

At least one person in your group use your phone's web browser to display: www.mrfelling.com/sa2

....

(the full activity is detailed in the filled-in notes if you were absent and need to refer to it)

....

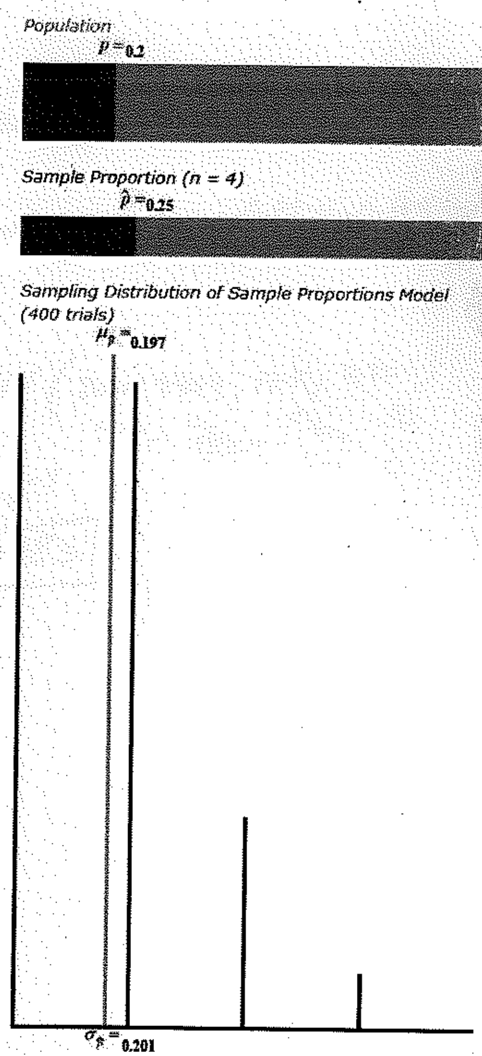
$$\boxed{\mu_{\hat{p}} = p \qquad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}}$$

Do these formulas correctly predict the mean and standard deviation for our sampling distribution of sample proportions in our phone applet?

$\mu_{\hat{p}} = 0.6$ (actual 0.607) ✓ $\qquad \sigma_{\hat{p}} = \sqrt{\frac{(0.6)(0.4)}{16}} = .122$ (actual 0.12) ✓

Is the shape of the sampling distribution of sampling proportions always Nearly Normal?
Let's investigate...if we 'reset' and select p=0.2 and n=4, then run many trials, we'll get something like this:

Not Nearly Normal.



Reset, and try the following settings:

p = 0.2, n = 8

p = 0.2, n = 16

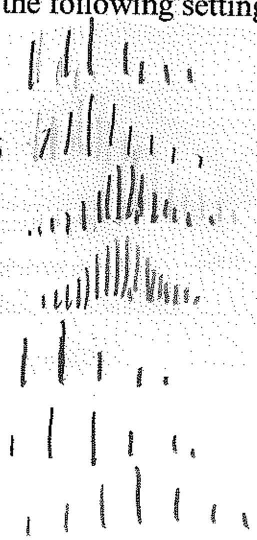
p = 0.2, n = 25

p = 0.2, n = 50

p = 0.2, n = 6

p = 0.3, n = 6

p = 0.5, n = 6



Write a few sentences describing the trends that you see:

If p is close to 0.5, sampling distribution is always approx Normal. But if p is not close to 0.5, distrib. is only approx normal for high n ($n>25$)

Normal if $np \geq 10$ and $nq \geq 10$

Practice Problems…

#1. Suppose that the mean adult weight is 175 pounds with a standard deviation of 25 pounds. An elevator in a building has a weight limit of 10 persons or 2000 pounds. What's the probability that the 10 people who get on the elevator will overload its weight limit?

- proportion or mean? "mean adult weight", weight is numerical → mean
- we need μ, σ for sampling distribution
- we are finding the probability that the mean weight of 10 people is more than $\frac{2000}{10} = 200$ lbs. (boundary for shading)
- population: $\mu = 175$ lbs, $\sigma = 25$ lb.

Sampling distribution model for means: $\mu_{\bar{x}} = \mu = 175$ lbs

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{25}{\sqrt{10}} = 7.906 \text{ lbs}$$

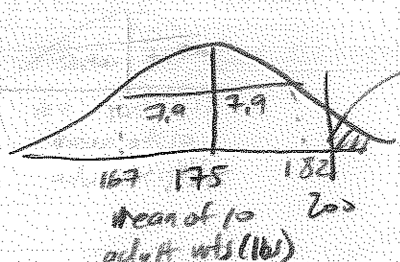
x check conditions

✓ – $n = 10$ (not ≥ 30 but we believe will be approx normal because we are sampling from a weight distribution which is likely to be normal)

✓ – $n < 10\%$? yes $10 < 10\%$ of all adults

✓ – independent? (we assuming there is no connection between the people on the elevator)

✓ – SRS? (assuming this sample is representative of the adult population)



$P(\bar{x} > 200) = \text{normalcdf}(\underset{\text{lower}}{200}, \underset{\text{upper}}{9999}, \underset{\text{mean}}{175}, \underset{\text{SD}}{7.906})$

$= 7.83 \times 10^{-4}$

$= \boxed{.000783}$

#2. Suppose that about 13% of the population is left-handed. A 200-seat school auditorium has been built with 15 "lefty seats", seats that have the built-in desk on the left rather than the right arm of the chair. In a class of 90 students, what's the probability that there will not be enough seats for the left-handed students?

- proportion or mean? (left-handed = Y, right-handed = N, categorical)
  "13%" proportion

- we need μ, σ for sampling distrib. of proportions
- we are finding the probability of not enough seats.
  This happens if there are more than 15 left handed students in our sample of 90 students.
  So we are finding the probability that the proportion of students who are left-handed is greater than $\frac{15}{90} = .167$ (boundary for stading)
- population: $p = .13$ (% left-handed)

For the sampling distribution of sample proportions:

$$\mu_{\hat{p}} = p = .13 \qquad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(.13)(.87)}{90}} = .0354$$

* check conditions

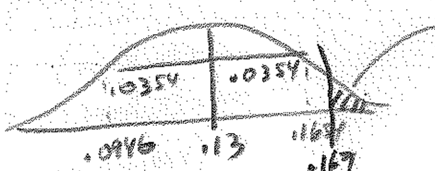
✓ – $np \geq 10$? $(90)(.13) = 11.7$
  $nq \geq 10$? $(90)(.87) = 78.3$

✓ – $n < 10\%$ pop? $90 < 10\%$ of all students

✓ – SRS? NO, but assume this class is representative of all classes.

$$P(\hat{p} > .167) = \text{normalcdf}(\underset{\text{lower}}{.167}, \underset{\text{upper}}{999}, \underset{\text{mean}}{.13}, \underset{\text{SD}}{.0354})$$

$$= \boxed{.1479}$$



proportion of sample of 90 which is left-handed

1. Coin tosses. In a large class of introductory Statistics students, the professor has each person toss a coin 16 times and calculate the proportion of his or her tosses that were heads. The students then report their results, and the professor plots a histogram of these several proportions.

a) What shape would you expect this histogram to be? Why?
b) Where do you expect the histogram to be centered?
c) How much variability would you expect among these proportions?
d) Explain why a Normal model should not be used here.

a) Symmetric, mound-shaped because it is a sampling distribution of proportions with $p = .5$

b) $\mu_{\hat{p}} = p = \boxed{0.5}$

c) $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.5)(.5)}{16}} = \boxed{.125}$

d) normal requires np & $nq \geq 10$
$np = (16)(.5) = 8$
sample size is not large enough

5. Just (un)lucky? One of the students in the introductory Statistics class in Exercise 1 claims to have tossed her coin 200 times and found only 42% heads. What do you think of this claim? Explain.

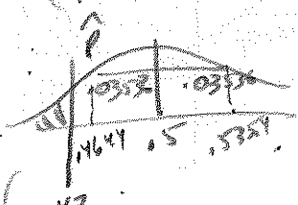
$\mu_{\hat{p}} = p = .5$

now $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.5)(.5)}{200}} = .03536$

two ways to analyze:

w/ probability:

$np = (200)(.5) = 100$
$nq = (200)(.5) = 100$ ≥ 10
(approx normal)



$P(\hat{p} < .42) = \text{normalcdf}(\underset{\text{lower}}{-999}, \underset{\text{upper}}{.42}, \underset{\text{mean}}{.5}, \underset{\text{SD}}{.03536})$
$= \boxed{.0118}$

Her claim is unlikely to be true... it would only happen randomly w/ probability .0118.

w/ z-score

$z = \frac{x - \mu}{\sigma} = \frac{.42 - .5}{.03536} = -2.26$

Her result is unlikely to be true because it is more than 2 standard deviations below the mean, and is unlikely to happen by chance.

16. Seeds. Information on a packet of seeds claims that the germination rate is 92%. What's the probability that more than 95% of the 160 seeds in the packet will germinate? Be sure to discuss your assumptions and check the conditions that support your model.

conditions

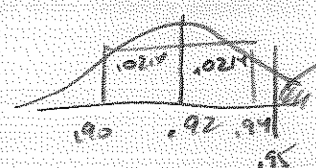
✓ — $np = (160)(.92) = 147$
$nq = (160)(.08) = 12.8 \geq 10$

✓ — 160 < 10% of all seeds

✓ — SRS (not sure, but assume a packet is representative)

$\mu_{\hat{p}} = p = .92$

$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.92)(.08)}{160}} = .0214$



$P(\hat{p} > .95) =$ normalcdf(.95, .999, .92, .0214)
lower, upper, mean, SD

$= \boxed{.08}$

17. Apples. When a truckload of apples arrives at a packing plant, a random sample of 150 is selected and examined for bruises, discoloration, and other defects. The whole truckload will be rejected if more than 5% of the sample is unsatisfactory. Suppose that in fact 8% of the apples on the truck do not meet the desired standard. What's the probability that the shipment will be accepted anyway?

$n = 150, \quad p = .08$

conditions

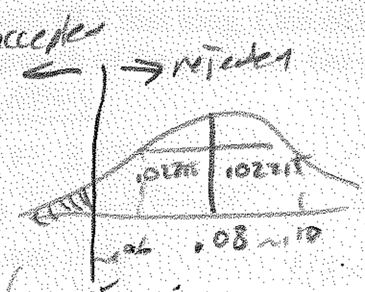
✓ — $np = (150)(.08) = 12$
$nq = (150)(.92) = 138 \geq 10$

✓ — SRS "random sample" stated

✓ — 150 < 10% of all apples

$\mu_{\hat{p}} = p = .08$

$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.08)(.92)}{150}} = .02215$



$P(\hat{p} < .05) =$ normalcdf(-999, .05, .08, .02215) $= \boxed{.0878}$
lower, upper, mean, SD

Chapter 18 Practice Quiz

AP Statistics Quiz B – Chapter 18 **Name________________**

1. **It is generally believed that electrical problems affect about 14% of new cars. An automobile mechanic conducts diagnostic tests on 128 new cars on the lot.**
   a. **Describe the sampling distribution for the sample proportion by naming the model and telling its mean and standard deviation. Justify your answer.**

condition
✓ — $np = (128)(.14) = 17.9$ ≥ 10
$nq = (128)(.86) = 110$
✓ — SRS (No, but assume representative of the population)
✓ — 128 < 10% of all cars

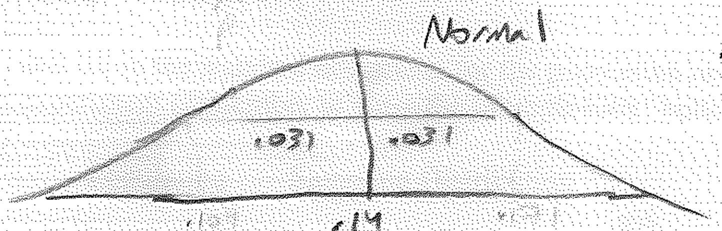
name $\boxed{\text{normal model}}$

$\mu_{\hat{p}} = p = \boxed{.14}$

$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.14)(.86)}{128}} = \boxed{.031}$

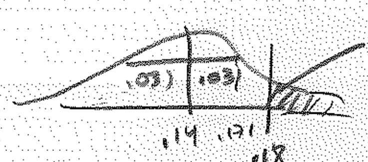
or can write $\boxed{N(.14, .031)}$

   b. **Sketch and clearly label the model.**



$\hat{p}$ (% of cars w/ electrical problems)

   c. **What is the probability that in this group over 18% of the new cars will be found to have electrical problems?**



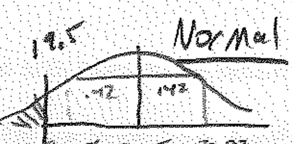
$P(\hat{p} > .18) = \text{normalcdf}(.18, 999, .14, .031)$ lower, upper, mean, SD
$= \boxed{.098}$

2. **Herpetologists (snake specialist) found that a certain species of reticulated python have an average length of 20.5 feet with a standard deviation of 2.3 feet. The scientists collect a random sample of 30 adult pythons and measure their lengths. In their sample the mean length was 19.5 feet long. One of the herpetologists fears that pollution might be affecting the natural growth of the pythons. Do you think this sample result is unusually small? Explain.**

Conditions
✓ — n = 30 ≥ 25
✓ — SRS "random sample"
✓ — Snakes indep (n < 10% pop)
30 < 10% of all snakes

$\mu_{\bar{x}} = \mu = 20.5$ ft

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.3}{\sqrt{30}} = .42$ ft



$P(\bar{x} < 19.5) = \text{normalcdf}(-999, 19.5, 20.5, .42)$ lower, upper, mean, SD
$= .0086$

yes, this is unusually small, because it should happen only .862 of the time by chance.

— or —

$z = \frac{19.5 - 20.5}{.42} = -2.38$

yes this is unusually small because it is more than 2 standard deviations below the average.