Lab 12: Chapters 8 and 9

1. A Washington Post-Schar School poll was conducted Oct. 1-9, 2019 among a random national sample of 1,007 U.S. adults. Results have a margin of error of ±3.5 percentage points.

   (a) In this poll, 60% of those surveyed said that President Trump does not uphold adequate standards for ethics in government. Exactly how many people surveyed said that President Trump does not uphold adequate standards for ethics in government?

   60% of 1007 = 0.6×1007 (a) 604

   (b) Interpret the meaning of the margin of error in the context of this problem. Assume a 95% confidence interval was used.

   It would be unusual for the sample percentage (60%) to differ from the actual percentage of U.S adults who said trump doesn't uphold adequate ethics standards by more that 3.5%.

2. A random sample of size 300 is to be selected from a population. Determine the mean and standard deviation of the sampling distribution of $\hat{p}$ for each of the following population proportions.

   $\sigma_{\hat{p}}$

   (a) $p = 0.20$
   (b) $p = 0.45$
   (c) $p = 0.70$
   (d) $p = 0.90$

   See attached

3. For which of the population proportions given in the previous exercise would the sampling distribution be approximately normal if $n = 40$? If $n = 75$?

   See attached

$x = 245$

$n = 935$

4. Suppose that 935 smokers each received a nicotine patch, which delivers nicotine to the bloodstream at a much slower rate than cigarettes do. Dosage was decreased to 0 over a 12-week period. Of these 935 people, 245 were still not smoking 6 months after treatment. Assume this sample is representative of all smokers.

(a) Use the given information to estimate the proportion of all smokers who, when given this treatment, would refrain from smoking for at least 6 months.

$$\hat{p} = \frac{x}{n} = \frac{245}{935} = 0.262$$

(a) 26.2%

(b) Verify that the conditions needed in order for the margin of error formula to be appropriate are met.

① The sample is rep. of the pop. ✓

② $n\hat{p} = 953(0.262) = 245$, $n(1-\hat{p}) = 935 - 245 = 708$ } both are ≥ 10, so the sampling dist is normal.

(c) Compute/find the value of the margin of error.

(c) 0.028 (2.8%)

$$M = 1.96\sigma_{\hat{p}} \doteq 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{0.262(1-0.262)}{935}}$$

(d) Interpret the meaning of the margin of error in the context of this problem.

It would be unusual for the estimate, 26.2%, to differ from the actual percentage of all smokers, who used the nicotine patch who were still not smoking 6 months after treatment by more than 2.8%.

(e) Construct a 95% confidence interval for the proportion of all smokers who, when given the nicotine patch, would refrain from smoking for at least 6 months.

(e) (0.234, 0.29)

① $\hat{p} - M = 0.262 - 0.028 = 0.234$

② $\hat{p} + M = 0.262 + 0.028 = 0.29$

(23.4%, 29%)

(f) Communicate the Result: Interpret the confidence interval.

We are 95% confident that the actual value of p is between 23.4% and 29%

(g) Communicate the Result: Interpret the confidence *level*.

A method has been used to produce a conf. interval that is successful in capturing the actual population percentage approximately 95% of the time.

$x = 504$
$n = 1007$

5. In an October survey on impeachment, 504 of 1007 randomly selected adult Americans reported that, in impeaching Trump, they think Democrats in Congress are distracting Congress from more important issues. Assume that this sample is representative of the population of U.S. adults.

(a) Use the given information to estimate the proportion of all adult Americans who think that, in impeaching Trump, Democrats in Congress are distracting Congress from more important issues.

$$\hat{p} = \frac{x}{n} = \frac{504}{1007}$$

(a) 0.50

(b) Verify that the conditions needed in order for the margin of error formula to be appropriate are met.

① The sample is rep. of the pop.

② There are 504 success and $1007 - 504 = 503$ failures in the sample, so both $n\hat{p}$ and $n(1-\hat{p}) \geq 10$.

(c) Compute/find the value of the margin of error. (Use a 99% confidence level)

$$M = 2.58\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2.58\sqrt{\frac{0.50(0.50)}{1007}}$$

(c) 0.041

(d) Interpret the meaning of the margin of error in the context of this problem.

It would be unusual for the estimate, 50%, to differ from the actual percentage of American adults with the belief by ~~more~~ more than 4.1%.

(e) Construct a 99% confidence interval for the proportion of all adult Americans who think that, in impeaching Trump, Democrats in Congress are distracting Congress from more important issues.

50% ± 4.1%

46% < p < 54.1%

(e) (46%, 54.1%)

(f) Communicate the Result: Interpret the confidence interval.

We are 99% confident that the actual percent of U.S. adults with the belief is between 46% and 54.1%

(g) Communicate the Result: Interpret the confidence *level*.

A method has been used to produce a confidence interval that is successful in capturing the actual pop. percentage approximately 99% of the time.

6. A random sample will be selected from the population of our college's students. The sample proportion $\hat{p}$ will be used to estimate p, the proportion of our college's students who work fulltime. For which of the following situations will the estimate tend to be closest to the actual value of p? Assume $n = 300$ in each situation.

$p = 0.80$ and

6. $p = 0.20$

The sampling dist that has the smallest margin of error gives the ~~best~~ closest estimate

5 cases

(a) $p = 0.80$
(b) $p = 0.60$
(c) $p = 0.50$
(d) $p = 0.40$
(e) $p = 0.20$

$$L2 = M = 1.96\sqrt{\frac{p(1-p)}{n}} = 1.96\sqrt{\frac{L1(1-L1)}{300}}$$

7. What pattern did you notice in the sequence of calculations you did to help you answer the previous questions?

The farther that p is from 50%, the better (closer) the est. is to the actual value

8. Consider taking a random sample from a population with $p = 0.25$.
   (a) What is the standard error of $\hat{p}$ for random samples of size 400?

$$\sigma_{\hat{p}} = \text{the st. dev of the sampling distn.} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.25(0.75)}{400}} = 0.022$$

   (b) Would the standard error of $\hat{p}$ be smaller for random samples of size 200 or samples of size 400?

$$\sigma_{\hat{p}} = \sqrt{\frac{0.25(0.75)}{200}} = 0.031$$

The st error is smaller for $n=400$.

   (c) Does cutting the sample size in half from 400 to 200 double the standard error of $\hat{p}$?

No, since 0.031 is not 2 times 0.022.

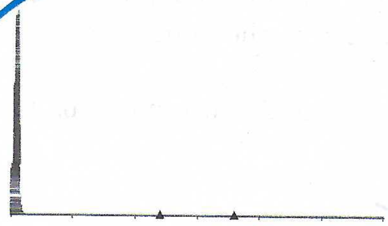


Figure 1: $p = 0.01$

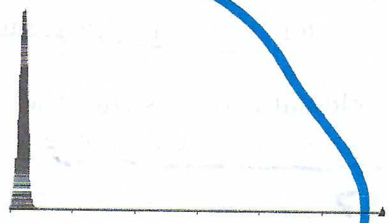


Figure 2: $p = 0.03$

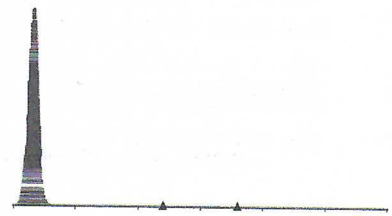


Figure 3: $p = 0.05$

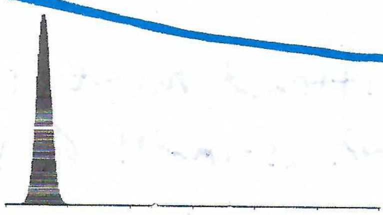


Figure 4: $p = 0.10$

Figure 5: $p = 0.20$

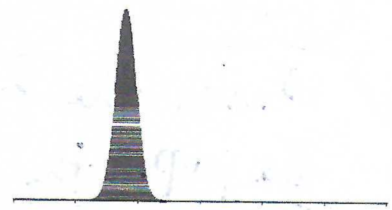


Figure 6: $p = 0.30$

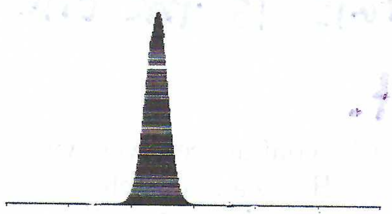


Figure 7: $p = 0.40$

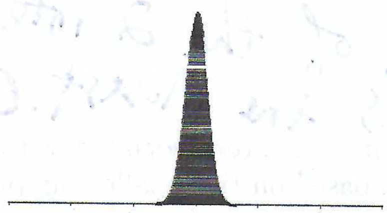


Figure 8: $p = 0.50$

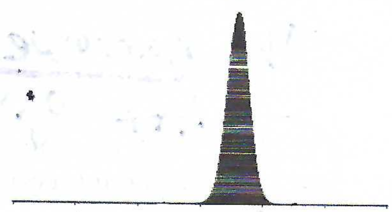


Figure 9: $p = 0.60$

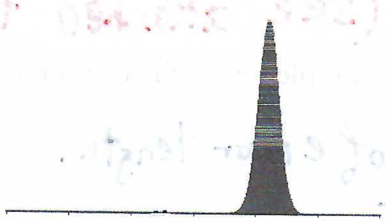


Figure 10: $p = 0.70$

Figure 11: $p = 0.80$

Figure 12: $p = 0.90$

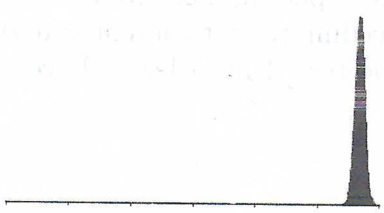


Figure 13: $p = 0.95$

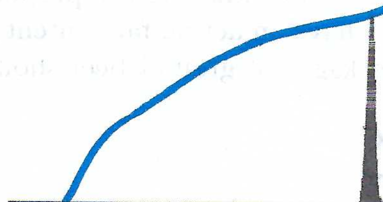


Figure 14: $p = 0.97$

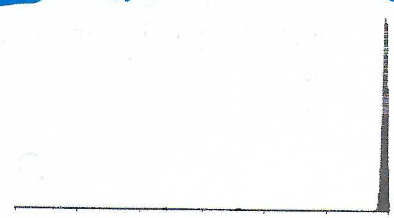


Figure 15: $p = 0.99$

For a fixed sample size, the standard error of $\hat{p}$ is greatest when $p = 0.50$. Therefore, $\hat{p}$ tends to produce more accurate estimates the farther the population proportion is from 0.50. For each graph below, $n = 300$.

When n is allowed to increase, and p is the same, the st dev gets smaller. This results in better estimates of p when sample size is relatively large

9. Suppose that county planners are interested in learning about the proportion of county residents who would pay a fee for a curbside recycling service if the county were to offer this service. Two different people independently selected random samples of county residents and used their sample data to construct the following confidence intervals for the proportion who would pay for curbside recycling:

$$Interval1 : (0.68, 0.74)$$

$$Interval2 : (0.68, 0.72)$$

(a) Explain how it is possible that the two confidence intervals are not centered in the same place.

Since the 2 samples included different members of the pop., the 2 samples gave different estimates of p.

(b) Which of the two intervals conveys more precise information about the value of the population proportion?

The narrowest of the 2 intervals is the one that gives the closest est.

(c) If both confidence intervals are associated with a 95% confidence level, which confidence interval was based on the smaller sample size? How can you tell?

The interval with the ~~widest~~ largest margin of error (Int. 1) has the smaller sample size. (see attached for more)

(d) If both confidence intervals were based on the same sample size, which interval has the higher confidence level? How can you tell?

The one with the largest margin of error length.

The widest interval, $\boxed{\text{Interval 1}}$

10. A consumer group is interested in estimating the proportion of packages of ground beef sold at a particular store that have an actual fat content exceeding the fat content stated on the label. How many packages of ground beef should be tested in order to have a margin of error of 0.05?

385

11. USA Today (January 24, 2012) reported that ownership of tablet computers and e-readers is soaring. Suppose you want to estimate the proportion of students at your college who own at least one tablet or e-reader. What sample size would you use in order to estimate this proportion with a margin of error of 0.03?

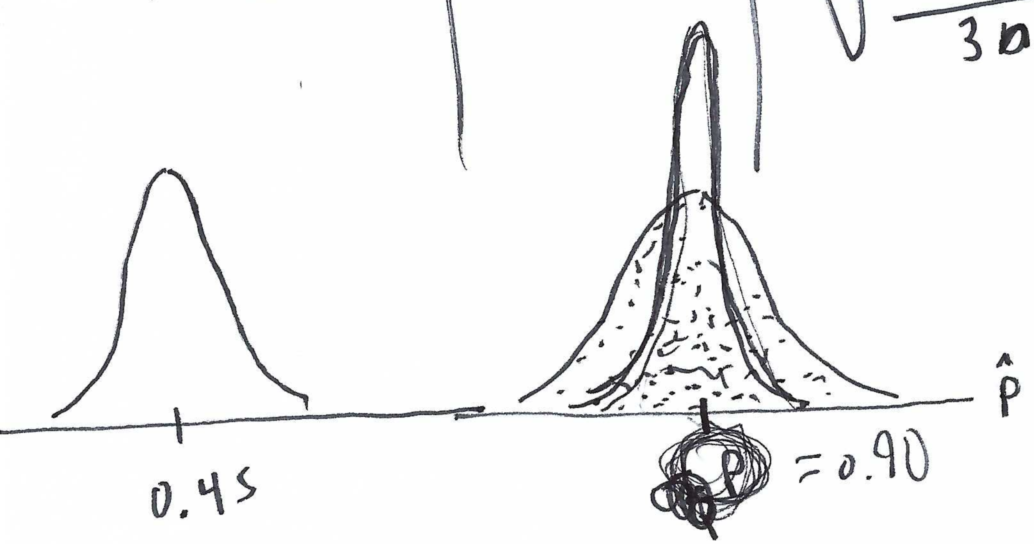
1068 (see attached)

(2)

$\boxed{n = 300}$

est. error $= |\hat{p} - p|$

p	$\mu_{\hat{p}}$	$\sigma_{\hat{p}}$
0.20	0.20	$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.2(1-0.2)}{300}} = \boxed{0.023 \ (2.3\%)}$
0.45	0.45	$\sqrt{\frac{0.45(1-0.45)}{300}} = \boxed{0.029 \ (2.9\%)}$
0.70	0.70	$\sqrt{\frac{0.7(1-0.7)}{300}} = \boxed{0.026 \ (2.6\%)}$
0.90	0.90	$\sqrt{\frac{0.9(1-0.9)}{300}} = \boxed{0.017 \ (1.7\%)}$



$$L2 = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{L1(1-L1)}{300}}$$

(3) Assume $n = 75$

p	np	$n(1-p)$	normal?
0.2	15	60	yes
0.45	33.75	41.25	yes
0.7	52.5	22.5	yes
0.9	67.5	7.5	no

3 Assume $n = 40$

p	successes np	failures $n(1-p)$	Normal?
0.2	8	32	no
0.45	18	22	yes
0.70	28	12	yes
0.90	36	4	no

$L2 = np = (40)(L1)$ $L3 = n(1-p) = 40(1-L1)$

Check: Are both np and $n(1-p)$ ≥ 10?

If so, we can assume the sampling dist. graph is normal.

n	$\sigma_{\hat{p}}$
L1	L2 $= \sqrt{\frac{0.5(1-0.5)}{L1}}$
30	
50	
100	
200	
300	

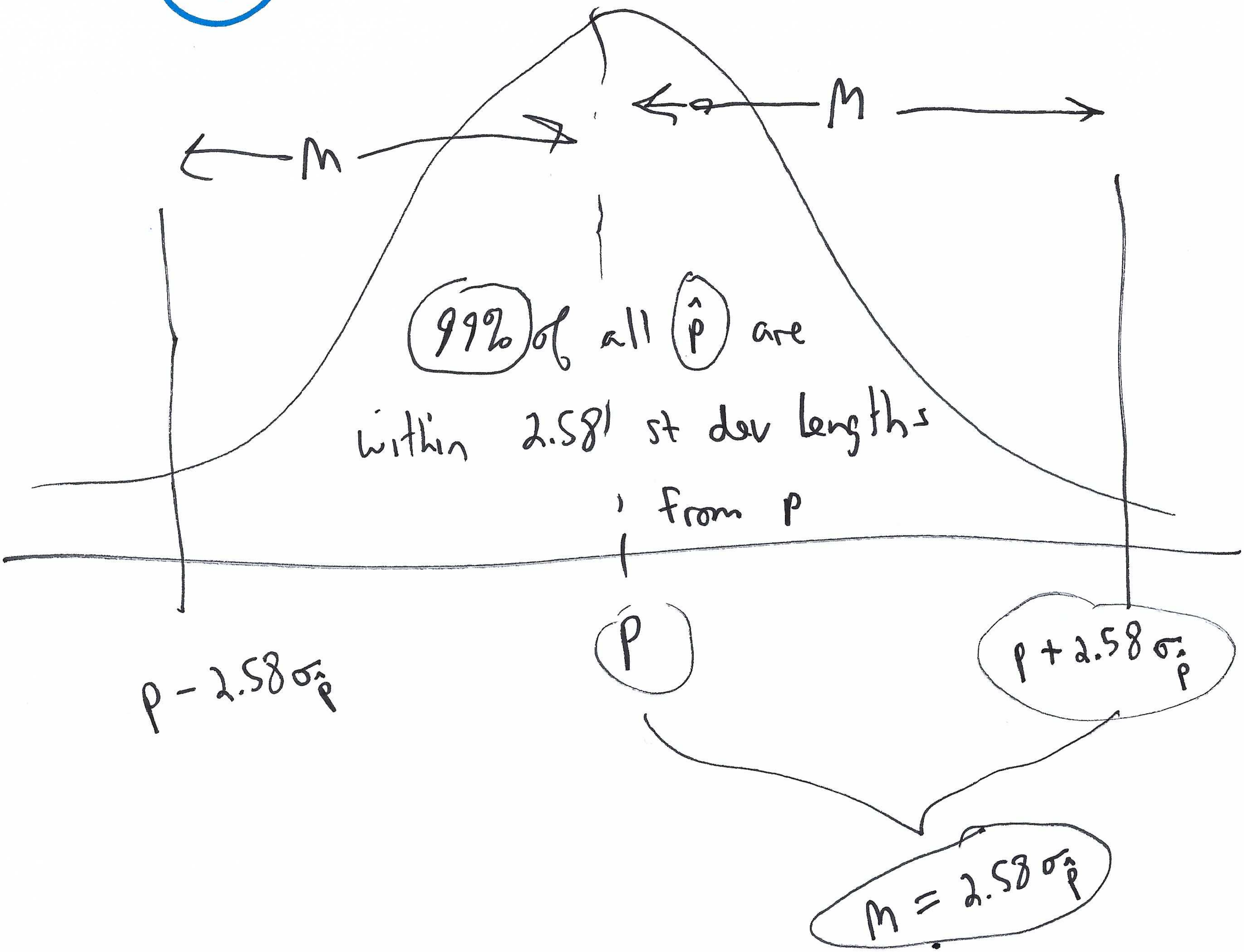
Assume $p=0.50$

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

$\boxed{\text{As } n \text{ increases, } \sigma_{\hat{p}} \text{ decreases}}$

As n decreases, $\sigma_{\hat{p}}$ decreases

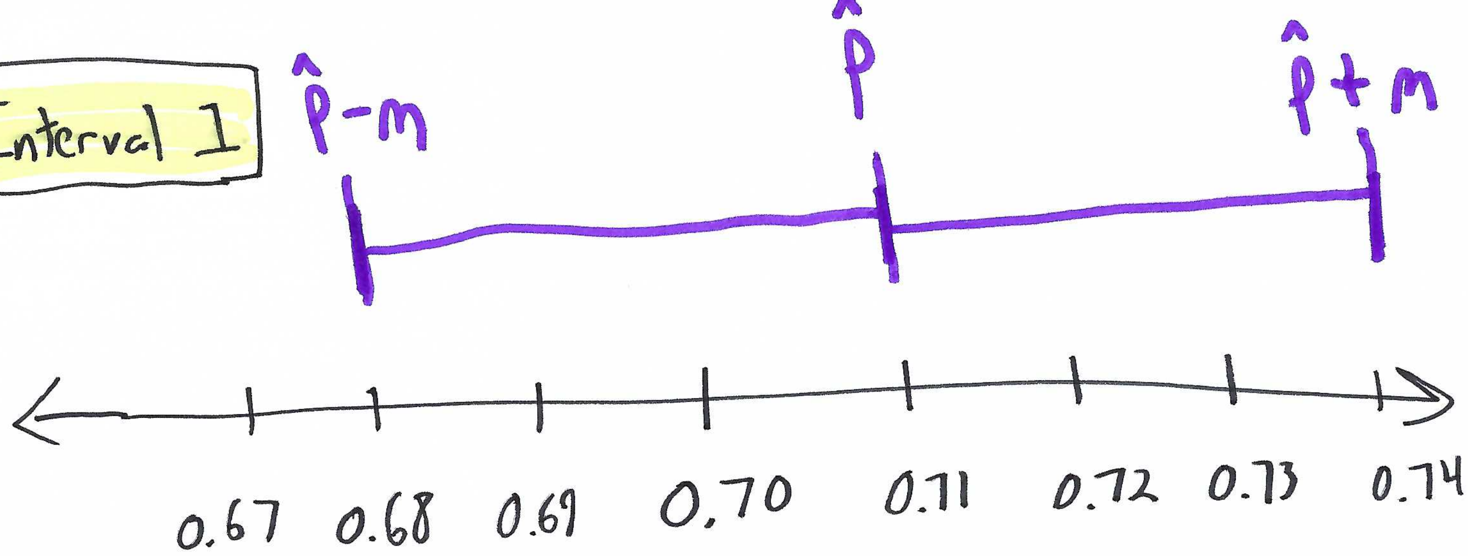
5



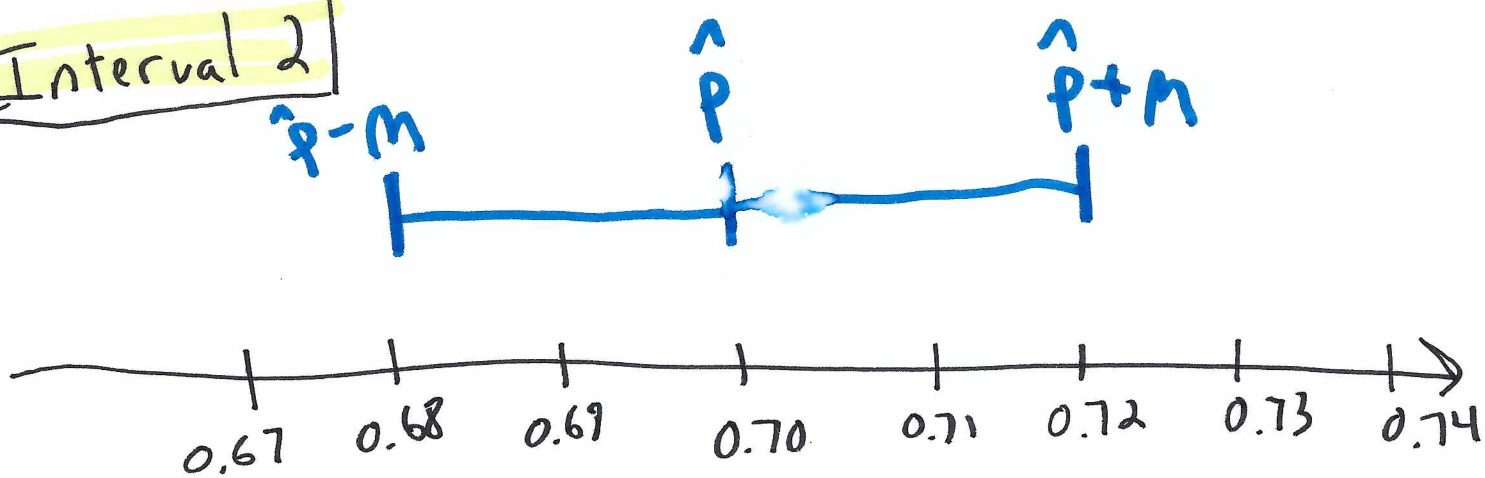
M
M
99% of all $\hat{p}$ are
within 2.581 st dev lengths
from p
p
$p - 2.58\sigma_{\hat{p}}$
$p + 2.58\sigma_{\hat{p}}$
$M = 2.58\sigma_{\hat{p}}$

9

Interval 1



Interval 2



9c

$$M = 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- as n increases, M decreases
- as n decreases, M increases

The larger the sample, the smaller the margin of error

(10)

$$n = p\cdot(1-p)\cdot\left(\frac{1.96}{M}\right)^2$$

$$=(0.5)(1-0.5)\left(\frac{1.96}{0.05}\right)^2$$

$$=(0.5)(0.5)\left(\frac{1.96}{0.05}\right)^2$$

$$=384.16$$

$$=\boxed{385}$$

If we take a random sample of 385 packages, our margin of error for estimating the actual percentage of packages that have an actual fat content exceeding the fat content stated on the label will be 5% or less.
This means our estimate for p from the sample will likely be within 5% of the actual value of p.

11

$$n = p \cdot (1-p) \cdot \left(\frac{1.96}{m}\right)^2$$

$$= (0.5)(1-0.5)\left(\frac{1.96}{0.03}\right)^2$$

$$= (0.5)(0.5)\left(\frac{1.96}{0.03}\right)^2$$

$$= 1067.\overline{1}$$

$$= 1068$$ Always round up to the <u>next</u> whole number

If we take a random sample of 1068 students, our margin of error for estimating p (the actual percent of students who have at least one tablet or e-reader) will be 3% or less.

This means our est. will likely be within 3% of the actual value of p.