Stat 13, Intro. to Statistical Methods for the Life and Health Sciences.

1. Power.
2. Confidence Intervals for a proportion and the dog sniffing cancer example.
3. CIs for a proportion and the Affordable Care Act example.
4. CIs for a population mean and the used cars example.

Start reading chapter 4.
http://www.stat.ucla.edu/~frederic/13/F16 .
HW2 is due Oct 18 and is problems 2.3.15, 3.3.18, and 4.1.23.

1. Power.

- Power is 1 – P(Type II error). Usually expressed as a function of μ.
- Recall Type I and Type II errors.
  - A type I error is a false positive. Rejecting the null when it is true.
  - A type II error is a false negative. Failing to reject the null when the null is false.

Power

- The probability of rejecting the null hypothesis when it is false is called the **power** of a test.
- Power is 1 minus the probability of type II error.
- We want a test with high power and this is aided by
  - A large effect size, i.e. true μ far from the parameter in the null hypothesis.
  - A large sample size.
  - A small standard deviation.
  - Significance level. A higher sign. level means greater power. The downside is that you increase the chance of making a type I error.

Estimation and confidence intervals.

Chapter 3

Chapter Overview

- So far, we can only say things like
  - "We have strong evidence that the long-run probability Buzz pushes the correct button is larger than 0.5."
  - "We do not have strong evidence kids have a preference between candy and a toy when trick-or-treating."
- We want a method that says
  - "I believe 68 to 75% of all elections can be correctly predicted by the competent face method."

Confidence Intervals

- Interval estimates of a population parameter are called **confidence intervals**.
- We will find confidence intervals three ways.
  - Through a series of tests of significance to see which proportions are plausible values for the parameter.
  - Using the standard deviation of the simulated null distribution to help us determine the width of the interval.
  - Through traditional theory-based methods.

Statistical Inference: Confidence Intervals

Section 3.1

Can Dogs Sniff Out Cancer?

Section 3.1

Can Dogs Sniff Out Cancer?

Sonoda et al. (2011). Marine, a dog originally trained for water rescues, was tested to see if she could detect if a patient had colorectal cancer by smelling a sample of their breath.

- She first smells a bag from a patient with colorectal cancer.
- Then she smells 5 other samples; 4 from normal patients and 1 from a person with colorectal cancer
- She is trained to sit next to the bag that matches the scent of the initial bag (the "cancer scent") by being rewarded with a tennis ball.

Can Dogs Sniff Out Cancer?

In Sonoda et al. (2011). Marine was tested in 33 trials.

- Null hypothesis: Marine is randomly guessing which bag is the cancer specimen (π = 0.20)
- Alternative hypothesis: Marine can detect cancer better than guessing (π > 0.20)

π represents her long-run probability of identifying the cancer specimen.

Can Dogs Sniff Out Cancer?

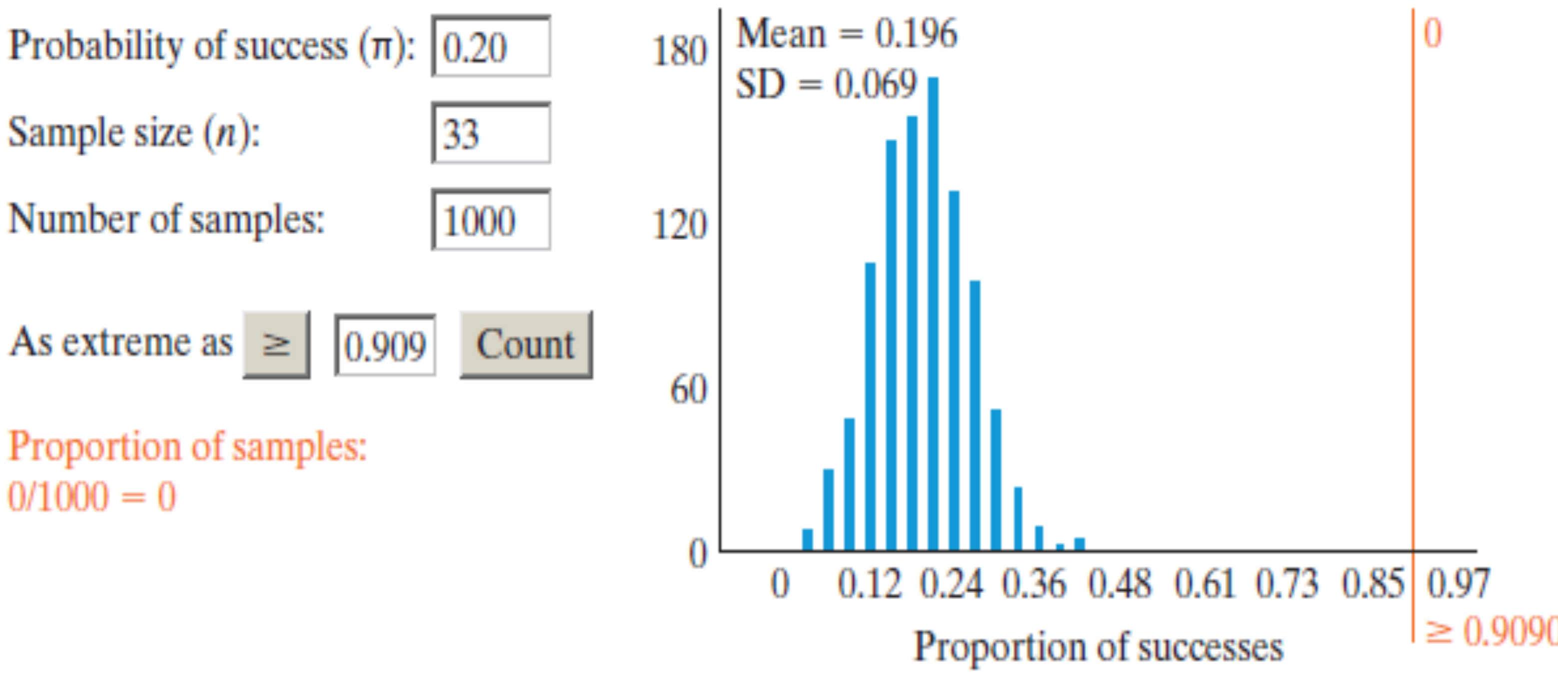
- 30 out of 33 trials resulted in Marine correctly identifying the bag from the cancer patient
- So our sample proportion is

$$\hat{p} = \frac{30}{33} \approx 0.909$$

- Do you think Marine can detect cancer?
- What sort of p-value will we get?

Can Dogs Sniff Out Cancer?

Our sample proportion lies more than 10 standard deviations above the mean and hence our p-value ~ 0.

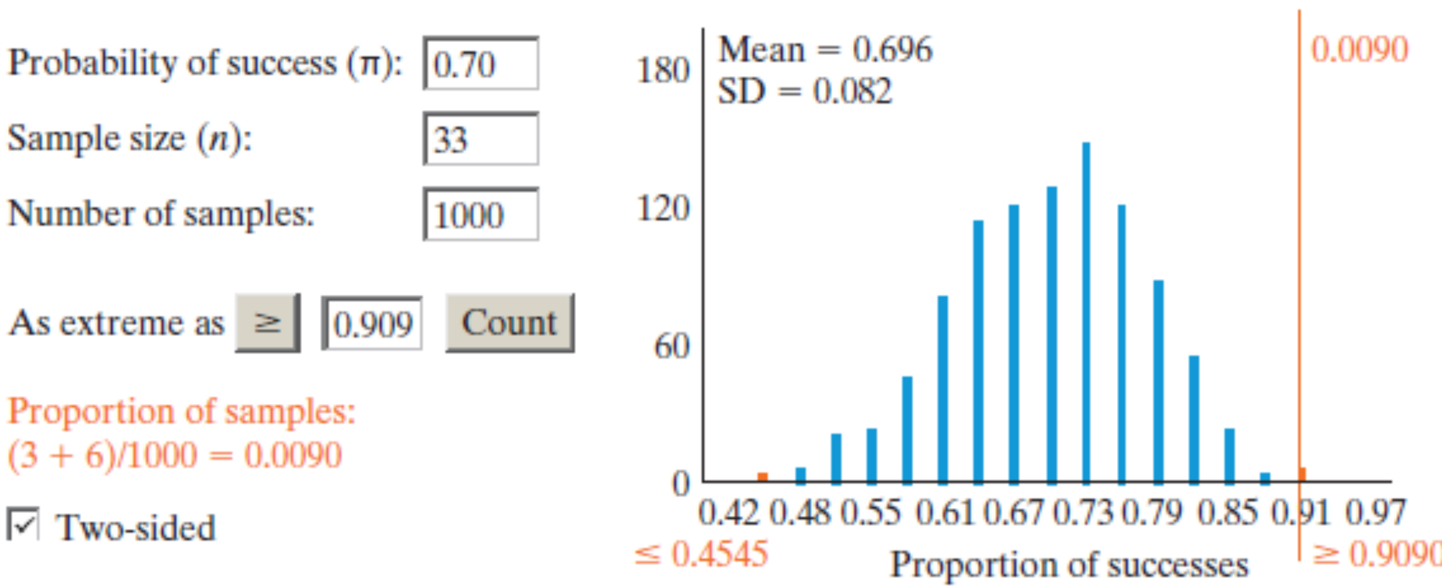


Can Dogs Sniff Out Cancer?

- Can we estimate Marine’s long run frequency of picking the correct specimen?
- Since our sample proportion is about 0.909, it is plausible that 0.909 is a value for this frequency. What about other values?
- Is it plausible that Marine’s frequency is actually 0.70 and she had a lucky day?
- Is a sample proportion of 0.909 unlikely if $\pi = 0.70$?

Can Dogs Sniff Out Cancer?

- H_0: $\pi = 0.70$ H_a: $\pi \neq 0.70$
- We get a small p-value (0.0090) so we can essentially rule out 0.70 as her long run frequency.

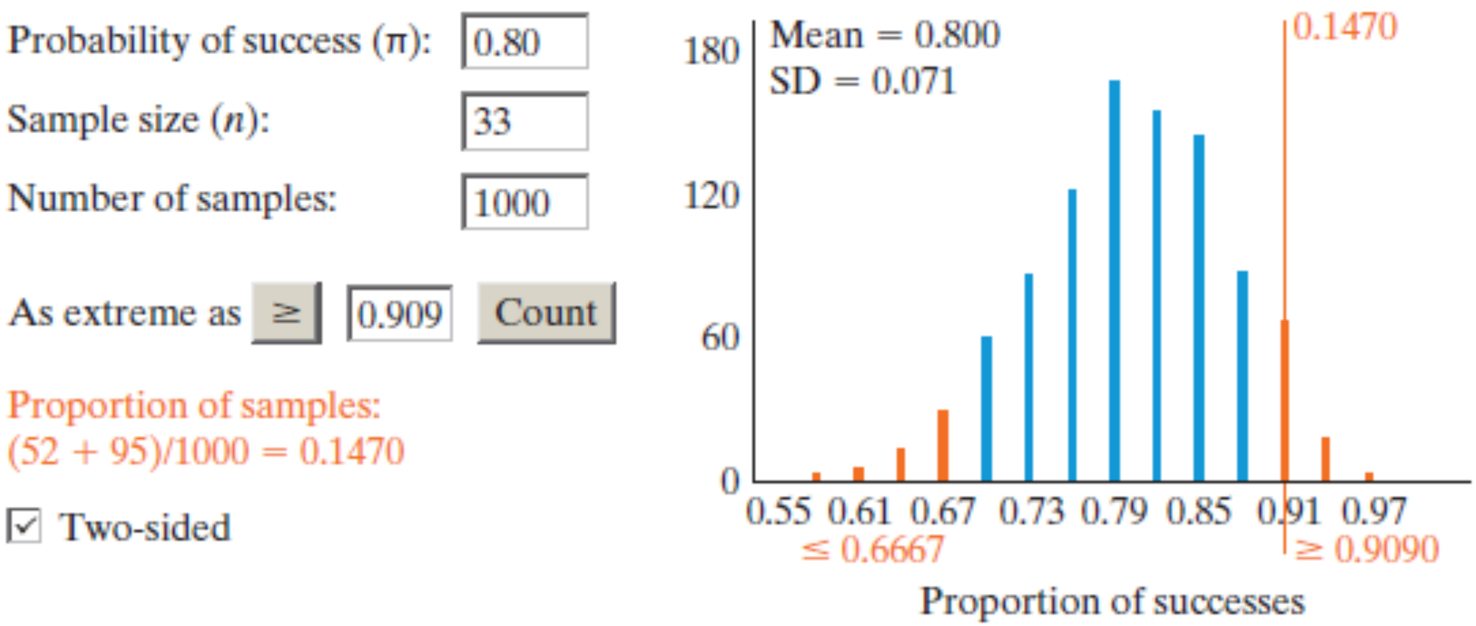


Can Dogs Sniff Out Cancer?

- What about 0.80?
- Is 0.909 unlikely if π = 0.80?

Can Dogs Sniff Out Cancer?

- H_0: $\pi = 0.80$ H_a: $\pi \neq 0.80$
- We get a large p-value (0.1470) so 0.80 is a ***plausible*** value for Marine's long-run frequency.



Developing a range of plausible values

- If we get a small p-value (like we did with 0.70) we will conclude that the value under the null is not plausible. This is when we reject the null hypothesis.
- If we get a large p-value (like we did with 0.80) we will conclude the value under the null is plausible. This is when we can't reject the null.

Developing a range of plausible values

- One could use software (like the one-proportion applet the book recommends) to find a range of plausible values for Marine’s long term probability of choosing the correct specimen.
- We will keep the sample proportion the same and change the possible values of π.
- We will use 0.05 as our cutoff value for if a p-value is small or large. (Recall that this is called the **significance level**.)

Can Dogs Sniff Out Cancer?

- It turns out values between 0.761 and 0.974 are plausible values for Marine's probability of picking the correct specimen.

Probability under null	0.759	0.760	0.761	0.762	…….	0.973	0.974	0.975	0.976
p-value	0.042	0.043	**0.063**	**0.063**		**0.059**	**0.054**	0.049	0.044
Plausible?	No	No	Yes	Yes	……… Yes	Yes	Yes	No	No

Can Dogs Sniff Out Cancer?

- (0.761, 0.974) is called a ***confidence interval***.
- Since we used 5% as our significance level, this is a 95% confidence interval. (100% – 5%)
- 95% is the ***confidence level*** associated with the interval of plausible values.

Can Dogs Sniff Out Cancer?

- We would say we are 95% confident that Marine's probability of correctly picking the bag with breath from the cancer patient from among 5 bags is between 0.761 and 0.974.
- This is a more precise statement than our initial significance test which concluded Marine's probability was more than 0.20.
- Sidenote: We do not say $P\{\pi \text{ is in } (.761, .974)\} = 95\%$, because π is not random. The *interval* is random, and would change with a different sample. If we calculate an interval this way, then $P(\text{interval contains } \pi) = 95\%$.

Confidence Level

- If we increase the confidence level from 95% to 99%, what will happen to the width of the confidence interval?

Can Dogs Sniff Out Cancer?

- Since the confidence level gives an indication of how sure we are that we captured the actual value of the parameter in our interval,
  to be more sure our interval should be wider.
- How would we obtain a wider interval of plausible values to represent a 99% confidence level?
  - Use a 1% significance level in the tests.
  - Values that correspond to 2-sided p-values larger than 0.01 should now be in our interval.

2SD and Theory-Based Confidence Intervals for a Single Proportion

Section 3.2

Introduction

- Section 3.1 found confidence intervals by doing repeated tests of significance (changing the value in the null hypothesis) to find a range of values that were plausible for the population parameter (long run probability or population proportion).
- This is a very tedious way to construct a confidence interval.
- We will now look at two others way to construct confidence intervals [2SD and Theory-Based].

The Affordable Care Act

Example 3.2

The Affordable Care Act

- A November 2013 Gallup poll based on a random sample of 1,034 adults asked whether the Affordable Care Act had affected the respondents or their family.
- 69% of the sample responded that the act had no effect. (This number went down to 59% in May 2014 and 54% in Oct 2014.)
- What can we say about the proportion of all adult Americans that would say the act had no effect?

The Affordable Care Act

- We could construct a confidence interval just like we did last time.
- We find we are 95% confident that the proportion of all adult Americans that felt unaffected by the ACA is between 0.661 and 0.717.

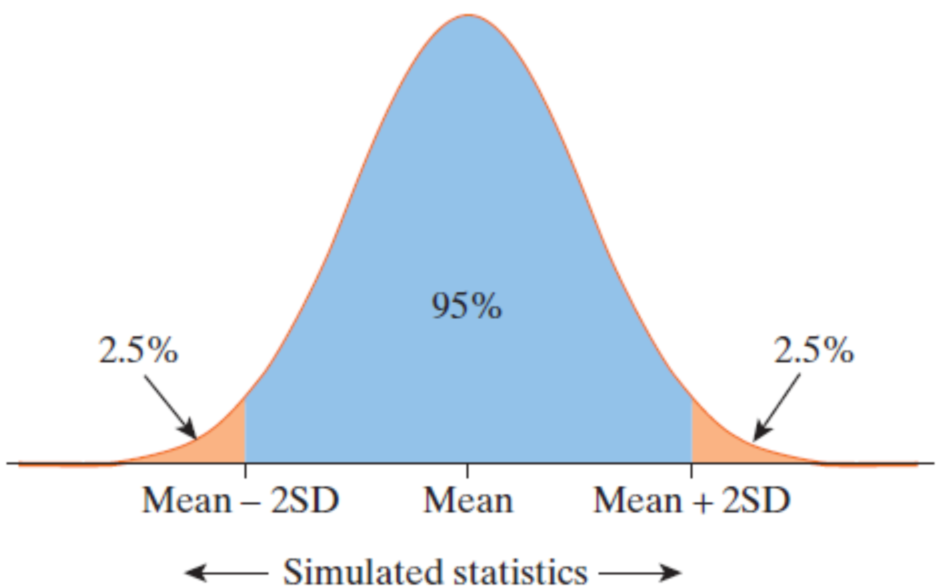
Probability under null	0.659	0.660	0.661	…………	0.717	0.718	0.719
Two-sided p-value	0.0388	0.0453	0.0514	…………	0.0517	0.0458	0.0365
Plausible value (0.05)?	No	No	Yes	…………	Yes	No	No

Short cut?

- The method we used last time to find our interval of plausible values for the parameter is tedious and time consuming.
- Might there be a short cut?
- Our sample proportion should be the middle of our confidence interval.
- We just need a way to find out how wide it should be.

2SD method

- When a statistic is normally distributed, about 95% of the values fall within 2 standard deviations of its mean with the other 5% outside this region



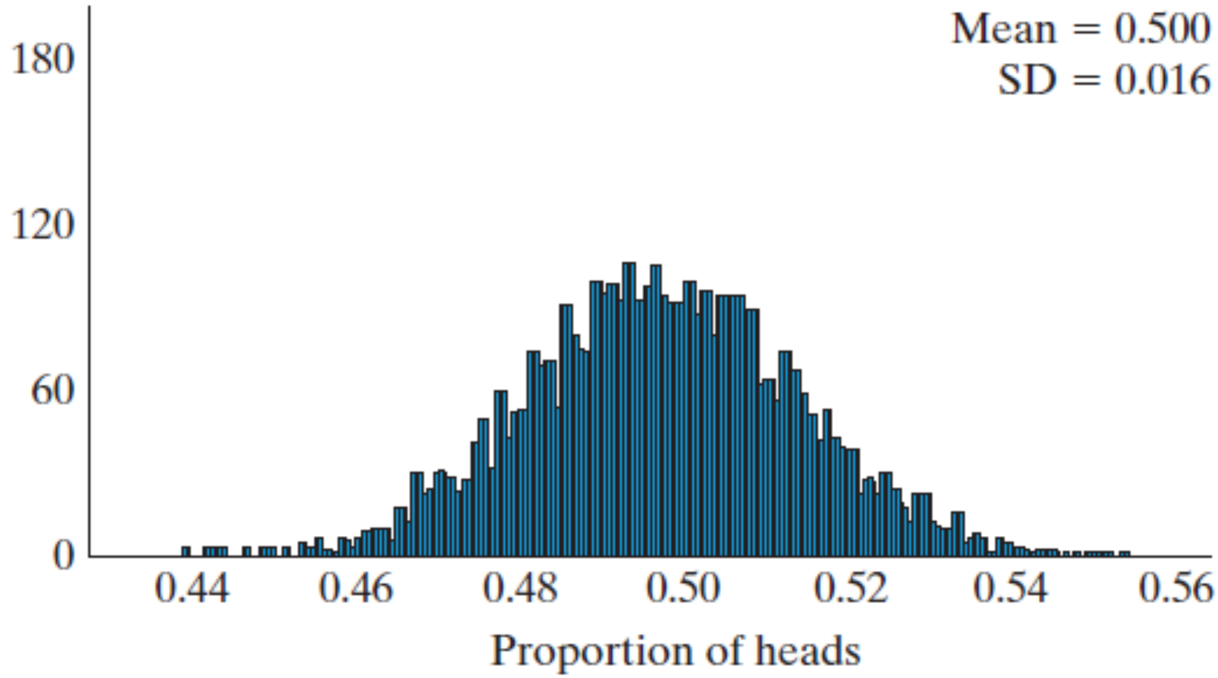
2SD method

- So we could say that a parameter value is plausible if it is within 2 standard deviations (SD) from our best estimate of the parameter, our observed sample statistic.
- This gives us the simple formula for a 95% confidence interval of

$$\hat{\boldsymbol{p}} \pm \boldsymbol{2SD}$$

Where do we get the SD?

- Null distribution for ACA with $\pi = 0.5$.



2SD method

- Using the 2SD method on our ACA data we get a 95% confidence interval

$$0.69 \pm 2(0.016)$$
$$0.69 \pm 0.032$$

- The $\pm$ part, like 0.032 in the above, is called the **margin of error.**
- The interval can also be written as we did before using just the endpoints; (0.658, 0.722)
- This is approximately what we got with our range of plausible values method (a bit wider).

Theory-Based Methods

- The 2SD method only gives us a 95% confidence interval
- If we want a different level of confidence, we can use the range of plausible values (hard) or theory-based methods (easy).
- The theory-based method is valid provided there are at least 10 successes and 10 failures in your sample.

Theory-Based Methods

- With theory-based methods we use normal distributions to approximate our simulated null distributions.
- Therefore we can develop a formula for confidence intervals.

$$\widehat{p} \pm multiplier \times \sqrt{\hat{p}(1-\hat{p})/n}.$$

For a 95% CI, the book suggests a multiplier of 2. Actually people use 1.96, not 2.

qnorm(.975) = 1.96.

qnorm(.995) = 2.58.

- Let’s check out this example using the theory-based method.
- Remember 69% of 1034 respondents were not affected.

$\widehat{p} \pm \textit{multiplier} \times \sqrt{\hat{p}(1-\hat{p})/n}$

$= 69\% \pm 2 \text{ x } \sqrt{.69(1-.69)/1034}$

$= 69\% \pm 2.88\%$.

With 1.96 instead of 2 it would be $69\% \pm 2.82\%$.

2SD and Theory-Based Confidence Intervals for a Single Mean

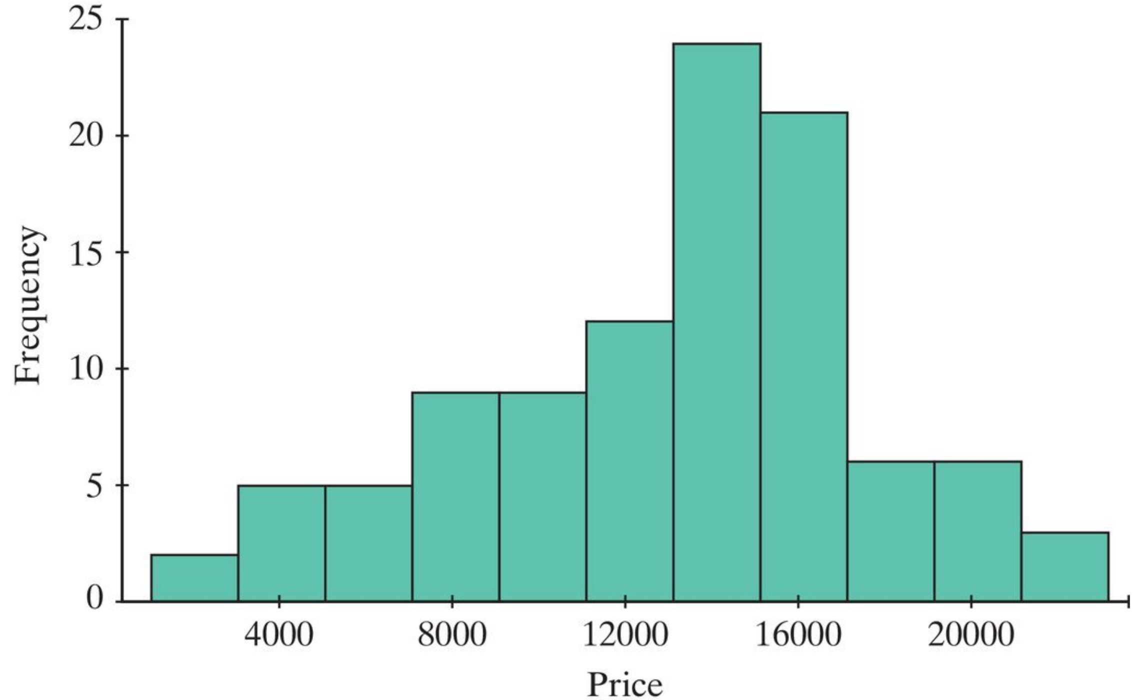
Section 3.3

Used Cars

Example 3.3

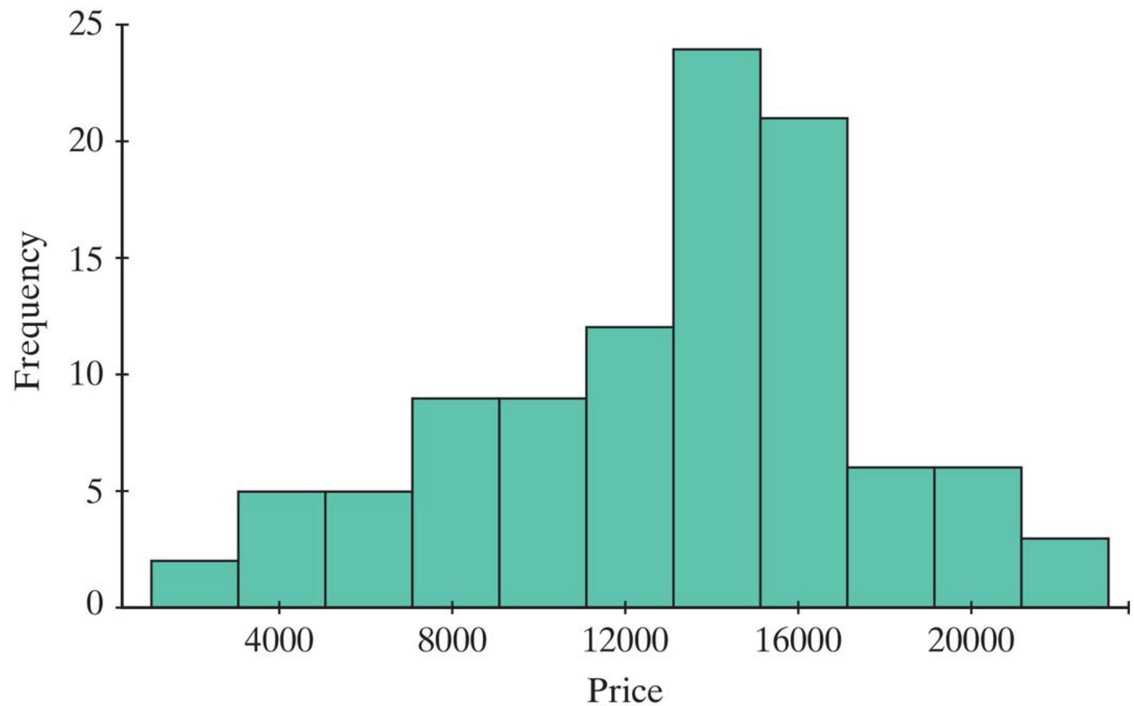
Used Cars

The following histogram displays data for the selling price of 102 Honda Civics that were listed for sale on the Internet in July 2006.



Used Cars

- The average of this sample is $\bar{x}$ = \$13,292 with a standard deviation of s = \$4,535.
- What can we say about μ, the average price of all used Honda Civics?



Used Cars

- While we should be cautious about our sample being representative of the population, let's treat it as such.
- μ might not equal \$13,292 (the sample mean), but it should be close.
- To determine how close, we can construct a confidence interval.

Confidence Intervals

- Remember the basic form of a confidence interval is:

  statistic ± multiplier × (SD of statistic)

  SD of statistic is also called Standard Error (SE).

- In our case, the statistic is $\bar{x}$ so we are write our 2SD confidence interval as:

$$\bar{x} \pm 2(\mathrm{SE})$$

Confidence Intervals

- It is important to note that the SD of $\bar{x}$ (the SE) and the SD of our sample (*s* = $4,535) are not the same.
- There is more variability in the data (the car-to-car variability) than in sample means.
- The SE is $s/\sqrt{n}$. Which means we can write a 2SD confidence interval as:

$$\bar{x} \pm 2\times\frac{s}{\sqrt{n}}$$

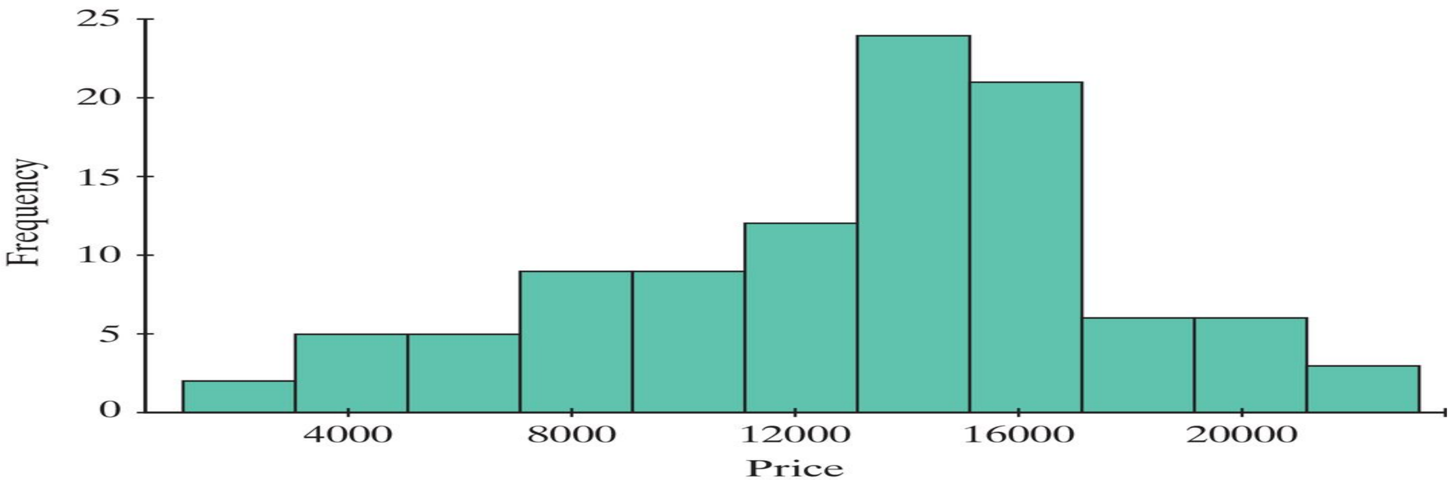
- This method will be valid when the null distribution is bell-shaped.

Summary Statistics

- A theory-based confidence interval is quite similar except it uses a multiplier that is based on a *t*-distribution and is dependent on the sample size and confidence level.
- For theory-based confidence interval for a population mean (called a one-sample t-interval) to be valid, the observations should be (approx.) independent, and either the population should be normal or n should be large. Check the sample distribution for skew and asymetry.

Confidence Intervals

- We find our 95% CI for the mean price of all used Honda Civics is from \$12,401.20 to \$14,182.80.
- Notice that this is a much narrower range than the prices of all used Civics.
- For a 99% confidence interval, it would be wider. The multiplier would be 2.6 instead of 1.96.



Factors that Affect the Width of a Confidence Interval

Section 3.4

Factors Affecting Confidence Interval Widths

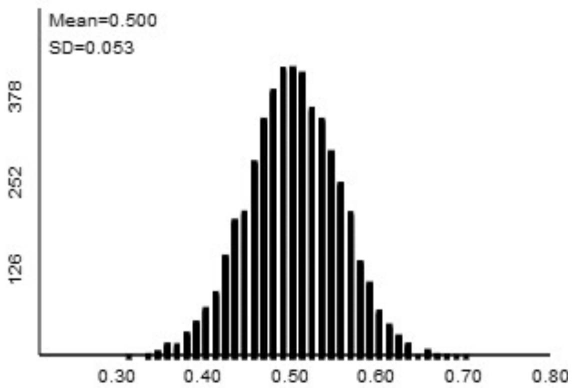
- **Level of confidence** (e.g., 90% vs. 95%)
  - As we increase the confidence level, we increase the width of the interval.
- **Sample size**
  - As sample size increases, variability decreases and hence the standard error will be smaller. This will result in a narrower interval.
- **Sample standard deviation**
  - A larger standard deviation, s, will yield a wider interval.
  - For sample proportions, wider intervals when $\hat{p}$ is closer to 0.5. s = √ [$\hat{p}$ (1-$\hat{p}$)].

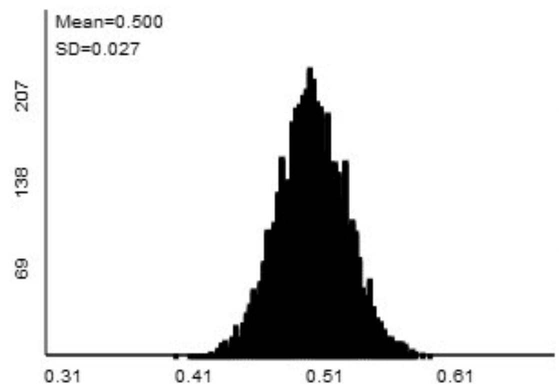
Level of Confidence

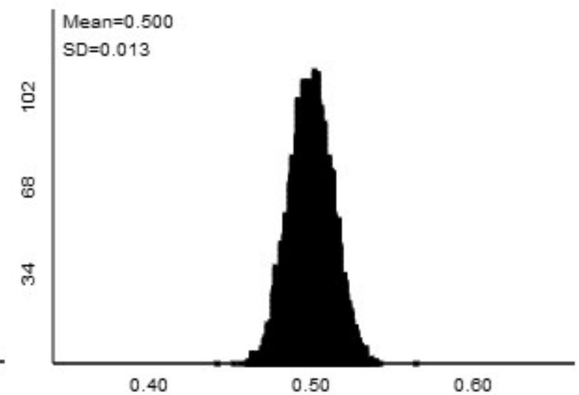
- If we have a wider interval, we should be more confident that we have captured the population proportion or population mean.
- We could see this with repeated tests of significance.
  - A higher confidence level corresponds to a lower significance level, and one must go farther to the left and farther to the right in our tables to get our confidence interval.

Sample Size

- We know as sample size increases, the variability (and thus standard deviation) in our null distribution decreases







n = 90 (SD = 0.054)   *n* = 361 (SD = 0.026)   *n* = 1444 (SD = 0.013)

Sample size	90	361	1444
SD of null distr.	0.053	0.027	0.013
Margin of error	2 x SD = 0.106	2 × SD = 0.054	2 × SD = 0.026
Confidence interval	(0.091, 0.303)	(0.143, 0.251)	(0.171, 0.223)

Sample Size

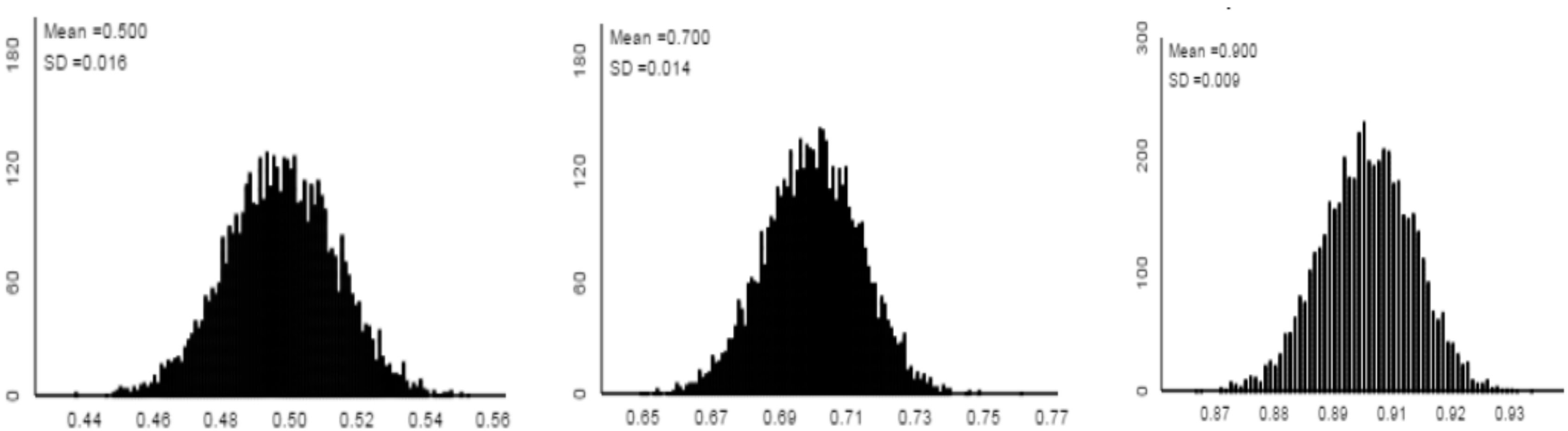
- (With everything else staying the same) increasing the sample size will make a confidence interval narrower.

Notice:

- The observed sample proportion is the midpoint. (that won't change)
- Margin of error is a multiple of the standard deviation so as the standard deviation decreases, so will the margin of error.

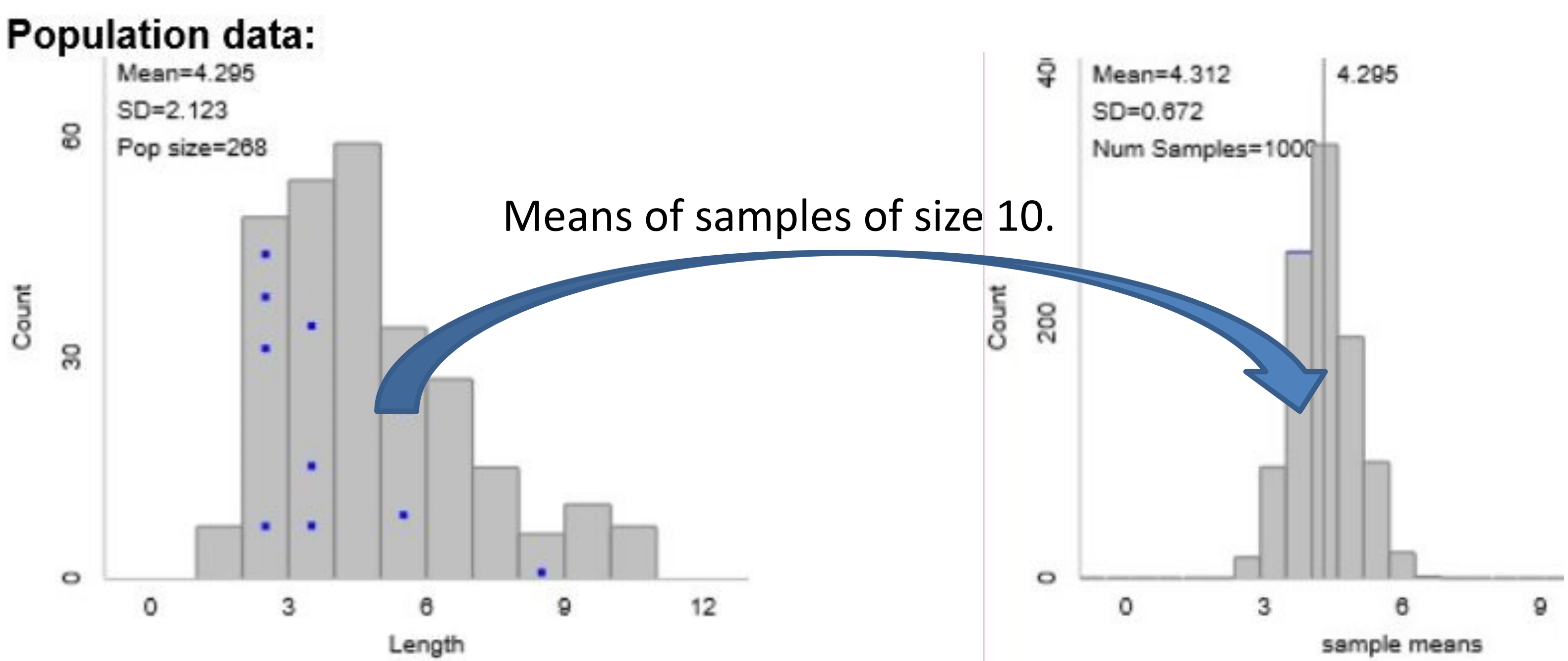
Value of $\hat{p}$
(or the value used for π under the null)

- As the value that is used under the null gets farther away from 0.5, the standard deviation of the null distribution decreases.
- When this standard deviation is used in the 2SD method, the interval gets gradually narrower.



Standard Deviation

- Suppose we are taking repeated samples of a population.
- How do we estimate what the standard deviation of the null distribution (standard error) will be? $s/\sqrt{n}$.



Standard Deviation

- The SD of the null distribution is approximated by $s/\sqrt{n}$.
- Remember that $2(s/\sqrt{n})$ is approximately the margin of error for a 95% confidence interval, so as the standard deviation of the sample data (s) increases so does the width of the confidence interval.
- Intuitively this should make sense, more variability in the data should be reflected by a wider confidence interval.

Formulas for Theory-Based Confidence Intervals

$$\hat{p} \pm multiplier \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad \bar{x} \pm multiplier \times \frac{s}{\sqrt{n}}$$

- The width of the confidence interval increases as level of confidence increases (multiplier)
- The width of the confidence interval decreases as the sample size increases
- The value $\hat{p}$ also has a more subtle effect. The farther it is from 0.5 the smaller the width.
- The width of the confidence interval increases as the sample standard deviation increases.

What does 95% confidence mean?

- If we repeatedly sampled from a population and constructed 95% confidence intervals, 95% of our intervals will contain the population parameter.
- Notice the interval is the random event here.

What does 95% confidence mean?

- Suppose a 95% confidence interval for a mean is 2.5 to 4.3. We would say we are 95% confident that the population mean is between 2.5 and 4.3.
  - Does that mean that 95% of the data fall between 2.5 and 4.3?
    - No
  - Does that mean that in repeated sampling, 95% of the sample means will fall between 2.5 and 4.3?
    - No
  - Does that mean that there is a 95% chance the population mean is between 2.5 and 4.3?
    - Not quite but close.

What does 95% confidence mean?

- What does it mean when we say we are 95% confident that the population mean is between 2.5 and 4.3?
  - It means that if we repeated this process (taking random samples of the same size from the same population and computing 95% confidence intervals for the population mean) repeatedly, 95% of the confidence intervals we find would contain the population mean.
  - P(confidence interval contains μ) = 95%.