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

1. CIs and dog sniffing cancer example, continued.
2. Sample size calculation.
3. 1.96SE and formula-based CIs for a proportion, ACA example.

Read chapter 4.

HW2 is due Wed, Feb12, 1159pm. 2.3.15, 3.3.18, and 4.1.23.
These problems are on the next 3 slides.

Midterm is Mon Feb24 in class.

On both Mon Feb3 and Wed Feb5, lecture will be recorded, rather than in person, and you will need to download it and watch it whenever you want. The links will be on the course website at recordedlectures.html .

The course website is http://www.stat.ucla.edu/~frederic/13/W25 .

Needles

Exercises 2.3.15 and 2.3.16 refer to the needle data.

2.3.15 Consider a manufacturing process that is producing hypodermic needles that will be used for blood donations. These needles need to have a diameter of 1.65 mm—too big and they would hurt the donor (even more than usual), too small and they would rupture the red blood cells, rendering the donated blood useless. Thus, the manufacturing process would have to be closely monitored to detect any significant departures from the desired diameter. During every shift, quality control personnel take a sample of several needles and measure their diameters. If they discover a problem, they will stop the manufacturing process until it is corrected.

a. Define the parameter of interest in the context of this study and assign an appropriate symbol to it.

b. State the appropriate null and alternative hypotheses using the symbol defined in (a).

c. Describe what a Type I error would be in this study. Also, describe the consequence of such an error in the context of this study.

d. Describe what a Type II error would be in this study. Also, describe the consequence of such an error in the context of this study.

3.3.18 Reconsider the investigation of the manufacturing process that is producing hypodermic needles. Using the data from the most recent sample of needles, a 90% confidence interval for the average diameter of needles is found to be (1.62 mm, 1.66 mm). For each of the following statements, say whether VALID or INVALID.

a. We are 90% confident that the average diameter of the sample of 35 needles is between 1.62 and 1.66 mm.

b. Based on the 90% confidence interval, there is evidence that the average diameter of needles produced by this manufacturing process is 1.65 mm.

c. Based on the 90% confidence interval, there is evidence that the average diameter of needles produced by this manufacturing process is different from 1.65 mm.

d. We are 90% confident that the average diameter of needles produced by this manufacturing process is between 1.62 and 1.66 mm.

e. About 90% of the needles produced by this manufacturing process have a diameter between 1.62 and 1.66 mm.

f. If we want to be more than 90% confident, we should take a larger sample of needles.

Colds and exercise

4.1.23 In November 2010, an article titled "Frequency of Colds Dramatically Cut with Regular Exercise" appeared in *Medical News Today*. The article was based on the findings of a study by researchers Nieman et al. (*British Journal of Sports Medicine*, 2010) that followed 1,002 people aged 18–85 years for 12 weeks, asking them to record their frequency of exercise (5 or more days a week? Yes or No) as well as incidences of upper respiratory tract infections (Cold during last week? Yes or No).

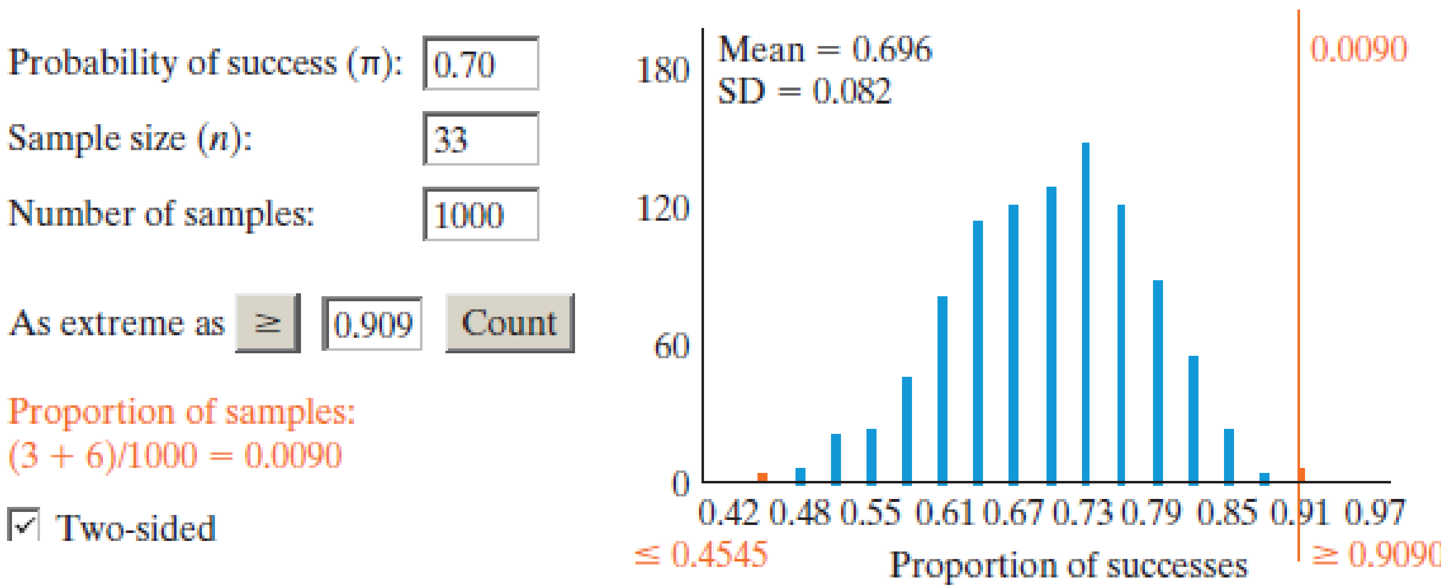
a. Identify the explanatory variable in this study. Also classify this variable as categorical or quantitative.

b. Identify the response variable in this study. Also classify this variable as categorical or quantitative.

c. Identify a confounding variable that provides an alternative explanation for the lower frequency of colds among those who exercised 5 or more days per week, compared to those who were largely sedentary.

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.

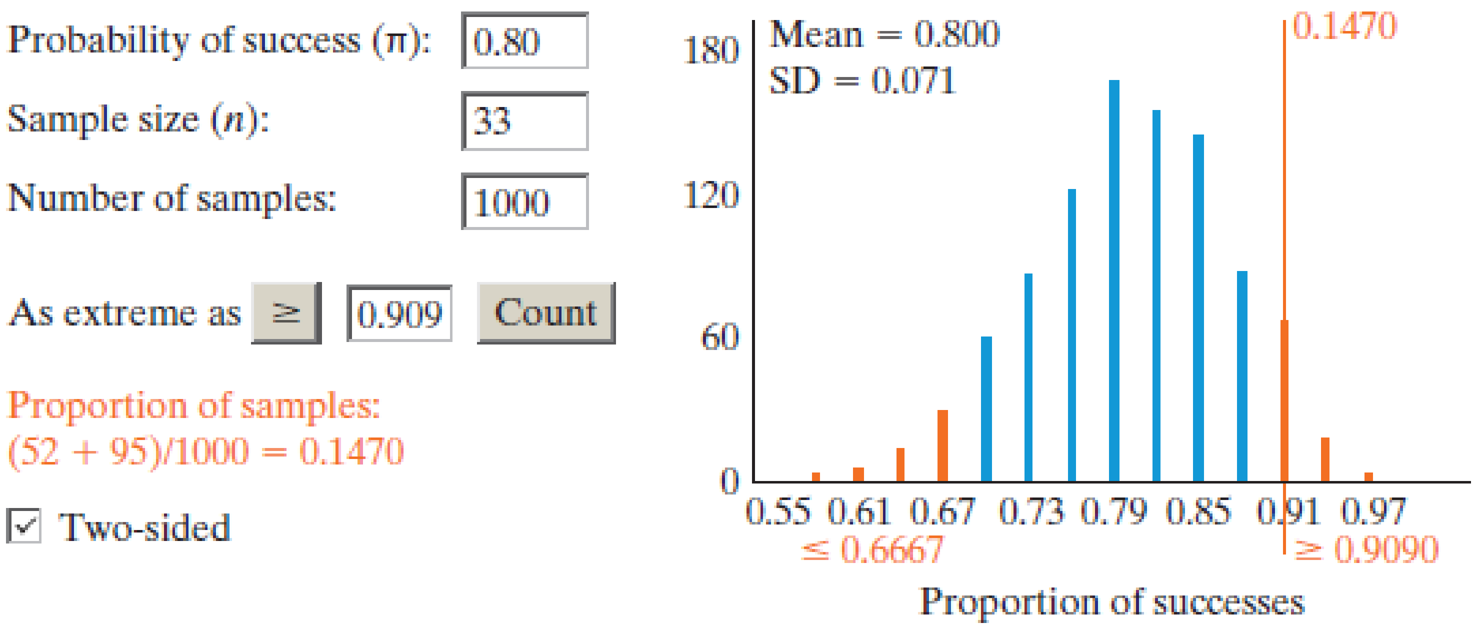


Can Dogs Sniff Out Cancer?

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

Can Dogs Sniff Out Cancer?

- H_0: π = 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{π 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(interval contains π) = 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.

2. Sample size calculation.

We previously saw that, when testing proportions, the standardized statistic Z = $\frac{\hat{p}-\pi}{SE}$,

where SE = $\sqrt{\pi(1-\pi)/n}$.

We also know that for the 2-sided Z-test, 1.96 is the cutoff for statistical significance.

If |Z| > 1.96, then p-value < 5%.

Suppose π = 50%, $\hat{p}$ = 70%, n = 10. How many more observations are needed to achieve statistical significance, if the effect size stays the same?

Z = $\frac{\hat{p}-\pi}{SE}$, SE = $\sqrt{\pi(1-\pi)/n}$.

We also know that for the 2-sided Z-test, 1.96 is the cutoff for statistical significance.

If |Z| > 1.96, then p-value < 5%.

Suppose π = 50%, $\hat{p}$ = 70%, n = 10. How many more observations are needed to achieve statistical significance, if the effect size stays the same?

We want to find n, so that 1.96 = $\frac{\hat{p}-\pi}{\sqrt{\pi(1-\pi)/n}} = \frac{.7-.5}{\sqrt{.5(1-.5)/n}}$

Squaring both sides, $1.96^2 = (.2)^2 / (.25/n) = .04n/.25$,

i.e. n = $.25 * 1.96^2 / .04 = 24.01$.

N needs to be *at least* 24.01, so really 25.

So we need 15 *more* observations.

3. 1.96SE and formula-based Confidence Intervals for a Single Proportion and ACA example.

Section 3.2

Introduction

- Previously we 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.
- This is a very tedious way to construct a confidence interval.
- We will now look at two others way to construct confidence intervals [1.96SE 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 get (0.661, 0.717).
- We are 95% confident that the proportion of all adult Americans that felt unaffected by the ACA is between 0.661 and 0.717.

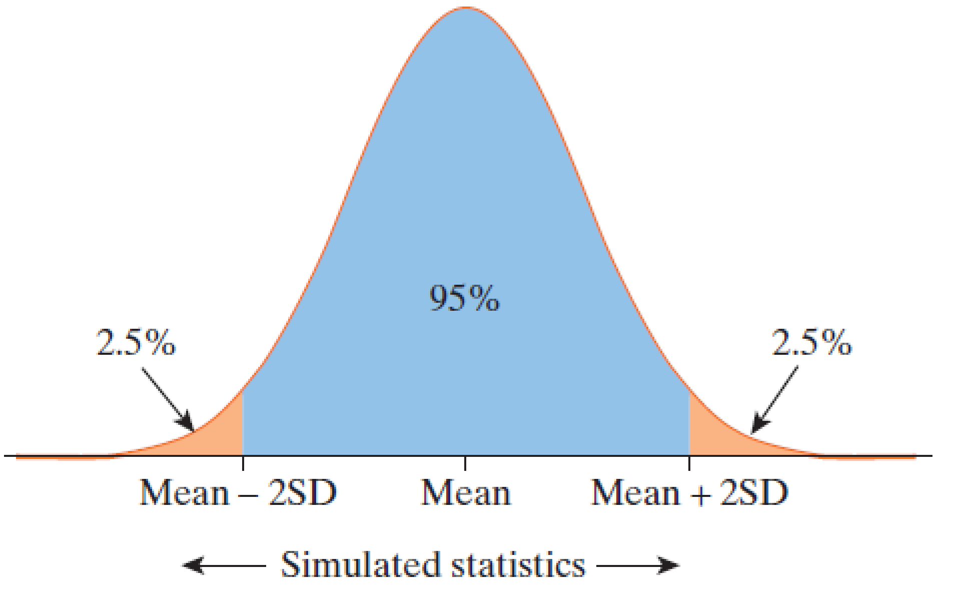
Probability under null	0.659	0.660	0.661	…………	0.717	0.718	0.719
Two-sided p-value	0.0388	0.0453	0.0514	…………	0.0517	0.0458	0.0365
Plausible value (0.05)?	No	No	Yes	…………	Yes	No	No

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.

1.96SE method

- When a statistic is normally distributed, about 95% of the values fall within 1.96 standard errors of its mean with the other 5% outside this region



1.96SE method

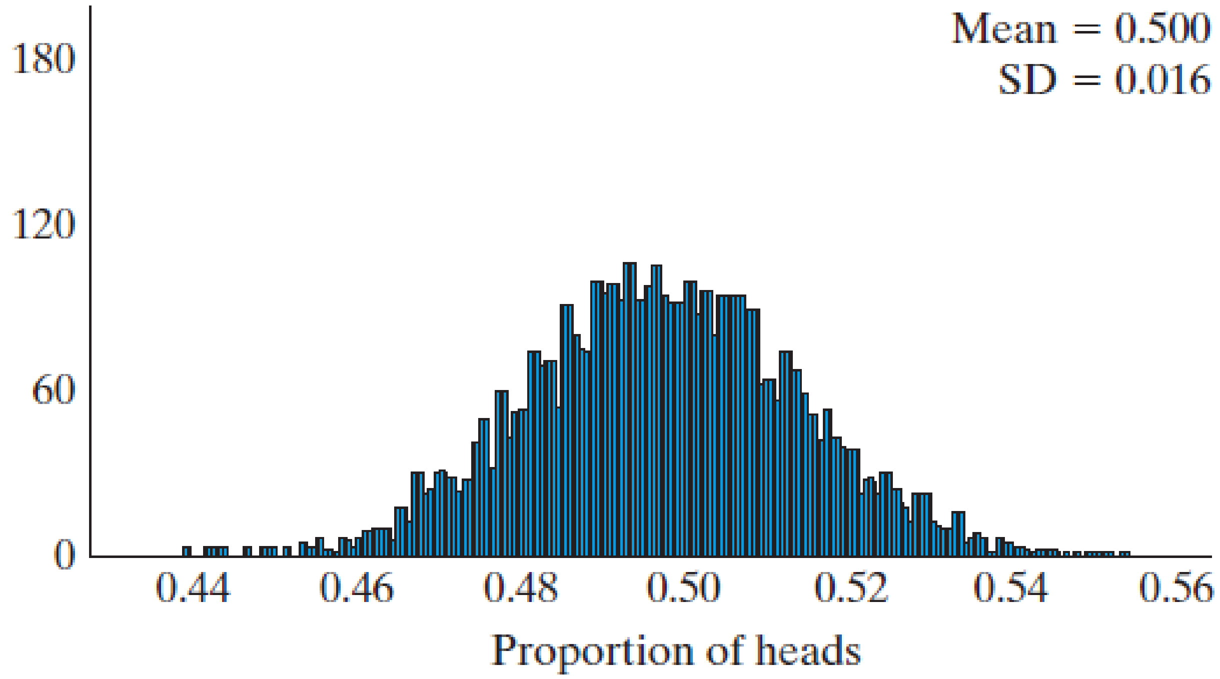
- So we could say that a parameter value is plausible if it is within 1.96 standard errors from our best estimate of the parameter, which is our observed sample statistic.
- This gives us the simple formula for a 95% confidence interval of

$$\boldsymbol{\hat{p} \pm 1.96SE}$$

Note that your book calls this the 2SD method but it really should be called the 1.96SE method.

Where do we get the SE?

- One way is via simulation.
- When the null hypothesis is $\pi = 0.5$, the SE = 0.016.



1.96SE method

- Using the 1.96SE method on our ACA data we get a 95% confidence interval

$$0.69 \pm 1.96(0.016)$$
$$0.69 \pm 0.031$$

- The $\pm$ part, like the 0.031 in the above, is called the **margin of error.**
- The interval can also be written as we did before using just the endpoints; (0.659, 0.721)
- This is approximately what we got using simulations, with our range of plausible values method. We had (0.661, 0.717).

Formula or Theory-Based Method

- The 1.96SE method is for 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 for CIs for a proportion, provided it's a Simple Random Sample (SRS) and there are at least 10 successes and 10 failures in your sample.

FORMULA FOR CIs FOR A PROPORTION.

- On the previous slides, we relied on simulations to tell us that the SE was 0.016. But we don't need this. In general for testing a proportion, under the null hypothesis, SE = $\sqrt{\pi(1-\pi)/n}$.
- For confidence intervals, we do not assume the null hypothesis, and since π is unknown, use $\hat{p}$ in its place:

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

For a 95% CI, the book suggests a multiplier of 2. Actually people use 1.96, not 2. This comes from a property of the normal distribution.

qnorm(.975) = 1.96.

qnorm(.995) = 2.58, the multiplier for a 99% CI.

- Going back to the ACA example, recall

69% of 1034 respondents were not affected.

With no default value of π, to get a 95% CI for $\widehat{p}$, use

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

$$= 69\% \pm 1.96 \times \sqrt{.69(1-.69)/1034}$$

$$= 69\% \pm 2.82\%.$$

With 2 instead of 1.96 it would be 69% $\pm$ 2.88%.

This is the formula we actually use for CIs for a proportion.

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

To review, the book first explains how to get a CI by repeated testing, then using the "2 SE" method where the SE is found via simulation, then gives you this formula. But the formula is actually the correct answer. The others are approximations and require simulation.