Section 5.2 Summary and Examples:

Confidence Interval for a Proportion

- The value in the center of the interval is the point estimate $\hat{p}$ which is the observed statistic from our actual study.
    - It represents our "best guess" for the parameter value p.
- The distance between the center of the interval and the lower/upper bounds is called the margin of error. It represents how uncertain we are about the parameter value.

When the following conditions are satisfied, the **95% confidence interval** is defined as

$$\boxed{\text{point estimate} \pm 1.96 \times \text{SE}}$$

- For 95% confidence, we can use a multiplier of 1.96.
- For other confidence levels (90%, 99%, etc.), we replace the 1.96 with a different multiplier value that depends on the confidence level.

Interpreting a 95% Confidence Interval

Typically we write the interpretation using one of two templates:

(1) We are 95% confident that the (name of the parameter you are estimating) is between (lower bound) and (upper bound).

(2) We estimate that the (name of the parameter you are estimating) is between (lower bound) and (upper bound), and we are 95% confident in our estimate.

Note: The true parameter value might be at the low end, at the high end, or somewhere in the middle – we don't know where exactly in the interval it is. "95% confident" says that in the long run, only 95% of these intervals will be correct.

What Does it Mean by "95% Confident"?

Suppose we took many samples and built a confidence interval from each sample using the equation "point estimate $\pm$ 1.96×SE" Then about 95% of those intervals would contain the true population proportion.

Worked with Saketh

Problem 1:
A survey of 1,000 adults asked "Do you favor or oppose 'sin taxes' on soda and junk food?" and 32% of them were in favor of taxing these foods.
(a) Find a 95% confidence interval for the proportion of adults favoring taxes on soda and junk food.

z = 1.96   p +- z*(((p(1-p))/n)^1/2

0.32 +1.96*(((0.32(1-0.32))/1000)^1/2 = 0.3489

0.32 -1.96*(((0.32(1-0.32))/1000)^1/2 = 0.2911

(0.2911, 0.3489)

(b) Interpret the confidence interval.

We are estimating the adult stance on sin taxes on soda and junk food, we are 95% confident that the adults favoring taxes on soda and junk food is between 29.11% and 34.89%

<u>Changing the Confidence Level</u>
In order to change the confidence level, we need to adjust the multiplier value. This multiplier is called a **critical value**.

Commonly Used Confidence Levels in Practice and Critical Values

Confidence Level	90%	95%	98%	99%
Critical Value: z^*	1.645	1.96	2.326	2.576

<u>One Sample Z Confidence Interval for a Proportion</u> (Large Sample)

$$\hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Required Condition: $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$

0.3020 - 2.576*(((0.3020(1-0.3020))/10)^1/2 = -0.0720

z = 2.576

k choose n * probablity^p * (1-probaily)^n-k

0.3020 +2.576*(((0.3020(1-0.3020))/10)^1/2 = 0.6760

10*0.8 > 10 = 8 > 10 so its not reliant

Problem 2:

p = 8/10 = 0.8

Heart Transplant Operations at St. George's Hospital

We use binomial formula instead

Data from St. George's Hospital: Of the last 10 heart transplants, 8 patients died within 30 days.

(a) Estimate, with a 99% confidence level, the true mortality rate at the hospital.

(10 choose 8) * (0.8)^8 * (0.2)^01-8 = 0.3020

(-0.0720, 0.6760)

p = 0.3020

Parameter: $p =$ true mortality rate at St. George's (actual probability of a death after a heart transplant operation at St. George's)

(b) The national mortality rate is 15% (30-day mortality rate for heart transplant patients). Do the data suggest that a heart transplant patient at St. George's has a higher probability of dying than the national rate? (If so, heart transplants will been suspended.)

0.15- 2.576*(((0.15(1-0.15))/10)^1/2 = 0.1409

0.15+ 2.576*(((0.15(1-0.15))/10)^1/2 = 0.4409

(-0.1409, 0.4409)

The data suggests that the upper and lower bounds are not in the national mortality rate, therfore st geroge's has a higher mortality rate than the national rate and heart transplants msut be suspended

Problem 3: We want to estimate the proportion of gun owners among the population of adults.

How many adults are needed to estimate it to have a 1% margin of error and a 99% confidence level? From a prior study, we learned that 34% of adults own a gun.

p +- z * SE

Margin of Error = z * ((p(1-p))/n)^1/2

(0.1/ 2.576) ^2 = (0.34(0.66)) / n

0.001506069 n =0.2244

2.576 *

1=2.576 *((0.34(0.66) / n)) ^1/2

n = 148.91

149 Adults

To Reduce the Margin of Error when Estimating p

We can reduce the margin of error (become more precise) by doing the following

- increase ________ the confidence level
- increase ________ the sample size n