**Activity 7.3.3 – Estimating Population Proportions**

In June 2015 Gallup – an influential national polling organization – reported that 52% of adults in the US have a great deal or quite a lot of confidence in the police. 52% ties the lowest this percentage has been in the past 22 years and reflects growing unease about policing in the United States. The sample survey was given to 1500 US adults in June 2015. The survey reported a margin of error of ±3.5 percentage points at a 95% confidence level. What does this mean?

The purpose of this survey is to estimate the population proportion of all US adults who have confidence in the police. The sample proportion (0.52) and margin of error (0.035) are combined to form an *interval estimate* for the population proportion. Interval estimates have the form:

The point estimate is the sample proportion; the margin of error is based on the sample proportion and sample size. In this example, the interval estimate is , which in interval form is (0.485, 0.555). Gallup concluded that, with 95% confidence, the proportion of US adults who have a great deal or quite a lot of confidence in the police is between 0.485 and 0.555. We will explore what “95% confidence” means by taking a look at how interval estimates are created.

**What is the Margin of Error?**

A sample proportion’s margin of error is the *likely* maximum possible distance between the sample proportion and population proportion. This distance is based on the sampling distribution. To see how this works, let’s examine a sampling distribution of sample proportions from a population with a *known* population proportion. We will find the margin of error for a 95% *confidence interval.*

A recent study found that 70% of teenagers aged 13 to 17 use smartphones. Let’s assume this is the true value of the population proportion. The dot plot below shows an empirical distribution of sample proportions from random samples of size *n* = 50, taken from a population with *p* = 0.70.

Samples = 500

Mean = 0.7

SE = 0.065



As we expect, the population proportion is the mean of the sample proportions. Based on the Empirical Rule, the interval that contains the *middle 95% of the sample proportions* is:

This interval is We can conclude that 95% of sample proportions are within 0.13 (in either direction) of the population proportion. 0.13 is the *margin of error* for a 95% confidence interval; it is the maximum possible distance between the population proportion and sample proportions in the middle 95% interval.

In general, the margin of error for a 95% confidence interval for a population proportion is two standard errors.

**How Is the Margin of Error Used?**

When the goal is to estimate the population proportion, the margin of error can be added to and subtracted from the sample proportion to form an interval estimate for the population proportion. Why? The key is to understand how sample proportions are distributed within a sampling distribution.

The dotplot below shows the middle 95% of sample proportions in a sampling distribution with

*p* = 0.70 and *n* = 50 . Note that the width of this interval is 0.26.



What happens if we create intervals of the *same width* centered at sample proportions instead of the population proportion *p*? The three number lines below show intervals of the same width (0.26) centered at three sample proportions, 0.65, 0.76, and 0.54.





The intervals centered at 0.65 and 0.76 (colored green) contain the population proportion *p* = 0.70. However the interval centered at 0.54 (colored red) does not. These examples highlight the reality that some intervals centered at sample proportions will contain the population proportion and others will not. If a sample proportion is in the middle 95%, its interval will contain the population proportion; if, however, a sample proportion lies outside of the middle 95%, its interval will not contain the population proportion. Since 95% of sample proportions generate intervals that *do* contain the population proportion, it is highly likely that a randomly selected sample proportion leads to an accurate interval estimate.

Of course, since the population proportion is unknown, we never know whether or not our interval estimate contains the population proportion.

**Approximating the Standard Error**

The standard error (standard deviation of sample proportions) is based on the population proportion *p*. Since our goal is to estimate this parameter, it is strange that we would use it to create the estimate. Since we typically don’t have an estimate for this parameter, we can instead calculate an *approximate standard error* by using the sample proportion to estimate the population proportion.

The following procedure outlines the process of constructing a 95% confidence interval for a population proportion based solely on sample statistics.

**Process for Constructing 95% Confidence Intervals for a Population Proportion**

1. Find a point estimate (sample proportion), .
2. Calculate the margin of error. The margin of error is two approximate standard errors.
3. Create the interval estimate:

Confidence intervals are interval estimates developed from sample data. The confidence level is the long-run probability that an interval estimate actually contains the true value of the population parameter, assuming this estimation process is repeated a large number of times. We will use the 95% confidence level but other confidence levels could be used.

1. In a random sample of 400 12th-grade students in Connecticut, 108 stated that the superpower they would most like to have is the ability to freeze time. Determine a 95% confidence interval for the proportion of all Connecticut 12th-grade students whose preferable super power is the ability to freeze time.
2. Determine the sample proportion.
3. Determine the approximate standard error.
4. Calculate the margin of error for a 95% confidence interval.
5. Create the 95% confidence interval and sketch the interval on the number line below.



1. Interpret the confidence interval.
2. In a random sample of 110 male 12th-grade students in Connecticut, 16 stated their favorite physical activity is basketball. Determine a 95% confidence interval for the proportion of all male 12th-grade students in Connecticut whose favorite physical activity is basketball.
3. Determine the sample proportion.
4. Determine the approximate standard error.
5. Calculate the margin of error for a 95% confidence interval.
6. Create the 95% confidence interval and sketch the interval on the number line below.



1. Interpret the confidence interval.
2. How would the confidence interval change if the sample size is doubled and the sample proportion remains the same?