A statistical hypothesis test is a process by which data is used to measure the amount of evidence for a carefully stated research question. The process of conducting a hypothesis test requires one to identify the parameter of interest, and in addition, provide a value for the parameter to be tested.

In contrast, a confidence interval is used to provide a likely range of values for a parameter. A confidence interval does not require pre-specification of a parameter. The goal of a confidence interval is to estimate the amount of inherent variation in a measurement over repeated sampling. The acceptable amount of inherent variation is called the margin-of-error.

Definitions

Confidence Interval: A likely range of values for a parameter

Margin-of-Error: The (acceptable) amount of inherent variation in a measurement over repeated sampling Always as a %

Example 5.1 Consider the following poll done by the Star Tribune, a newspaper from the Twin Cities. A total of 625 registered voters from Minnesota were surveyed. The survey question under consideration here is related to opinions regarding stricter gun laws in the United States.

StarTribune

Local Sports Business Opinion Variety Obituaries Classifieds Autos Housing Jobs

Minnesota Poll results: What Minnesotans think of current gun laws

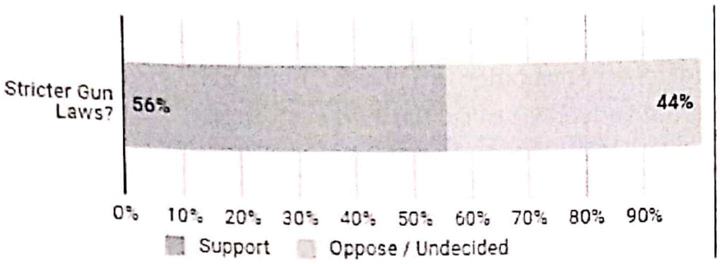
By Jeff Hargarten (http://www.startribune.com/jeff-hargarten/274254381/) , Dennis J. McGrath (http://www.startribune.com/dennis-j-mcgrath/10645391/) and Dave Braunger (http://www.startribune.com/dave-braunger/10644491/) Star Tribune

The Star Tribune Minnesota Poll interviewed 625 Minnesota registered voters between April 15 and April 18 about gun ownership, gun violence and gun control. Scroll down the page to see all breakdowns for each question, plus details about how the poll was conducted and the demographics of the respondents.

Survey Question: Do you support or oppose stricter gun laws in the United States?

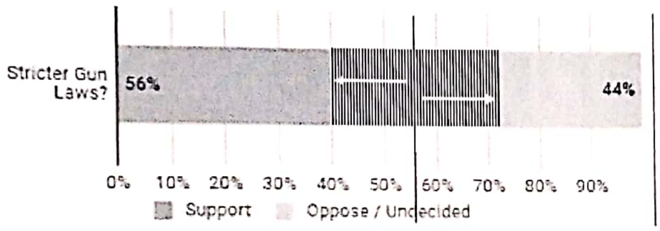
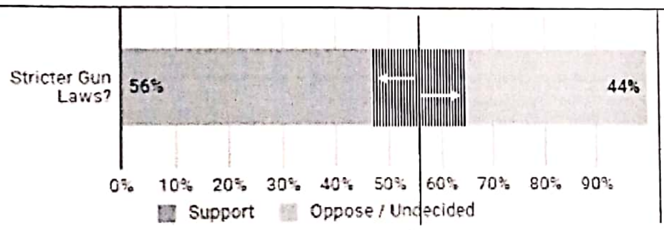
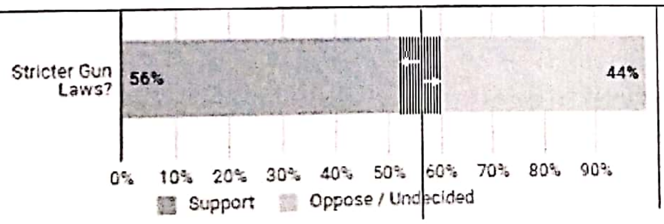
	Survey Response			Total
	Support	Oppose	Undecided	625
Stricter Gun Laws?	350 (56%)	194 (31%)	81 (13%)	

A graphic containing the survey responses – for simplicity the oppose / undecided were combined for the purposes of this example.



The following schematic shows the relationship between the confidence interval and margin-of-error. The confidence interval is shown with the striations and the up/down arrows show the positive/negative margin-of-error.

If the margin-of-error is large, then the range of likely values for the measurement of interest is large. This can be seen in the top graphic – there is a lot of inherent variation in the proportion of respondents who Support stricter gun control laws in the US in this graphic.

A lot of inherent variation	 56%	Confidence Interval: 40% up to 72% Margin-of-Error: $\pm$ 16%
Moderate amount of inherent variation	 56%	Confidence Interval: 48% up to 64% Margin-of-Error: $\pm$ 8%
Not much inherent variation	 56%	Confidence Interval: 52% up to 60% Margin-of-Error: $\pm$ 4%

Approach #1: Using a Simulation Model

A simulation model and its accompanying reference distribution can be used to estimate the amount of inherent variation in a measurement over repeated sampling. That is, a simulation model can be used to estimate the lower and upper endpoint of the confidence interval. From these endpoints, the margin-of-error can be computed.

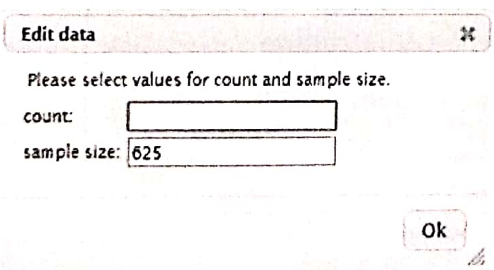
Necessary information for building a model	Stricter Gun Laws Case Study
• Number of trials, i.e. number of people who took poll	625
• The proportion that supported stricter gun laws	56%

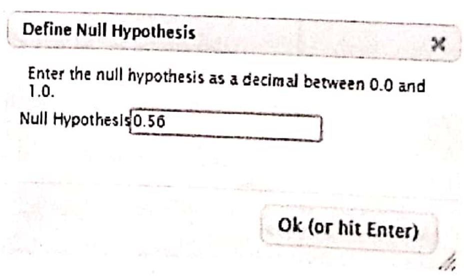
Note: Previously, a simulation model required the specification of a parameter. However, the purpose of a confidence interval is to obtain a likely range of values for the unknown parameter. Thus, the point estimate, i.e. the "best-guess", for the parameter will be used in the construction of the simulation model.

Quantity	Description
Scope-of-Inference	Minnesota Registered Voters
Parameter	π = the proportion of Minnesota Registered Voters who support stricter gun laws in the US
Study Participants	The 625 MN registered voters who participated in this survey
Statistic (or point estimate)	The measurement of interest, i.e. the proportion of survey respondents who support stricter gun laws in US. $\hat{\pi} = \frac{350}{625} = 56\%$

^ - coming from data
$\hat{\pi}$ - statistic
π - parameter

The simulation model will be setup using the outcomes from the survey. A total of 10,000 repeated simulations will be used to construct the reference distribution .



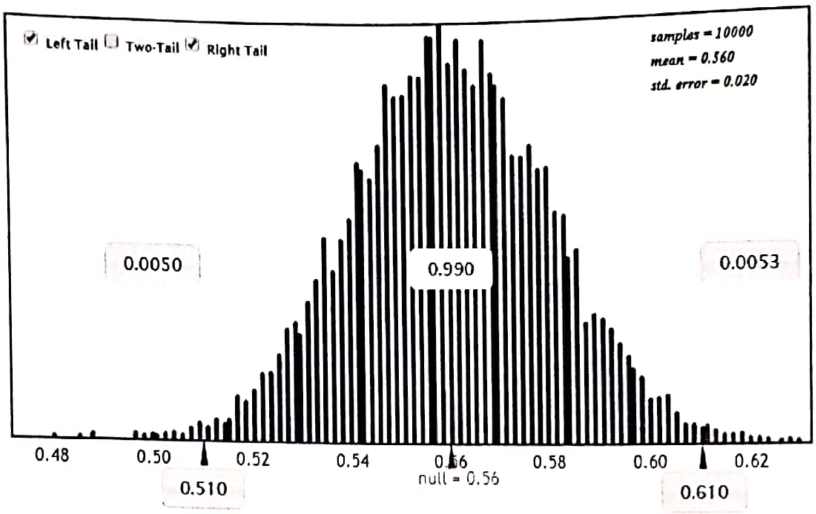


Finding the lower/upper endpoints of the confidence interval

A confidence interval is a range of likely value for a parameter over repeated sampling. The 5% rule is used to separate *likely* from *unlikely* values in a distribution. For simplicity, only two-sided confidence intervals will be considered here. Finally, confidence intervals and margin-of-errors are usually computed using the proportion (instead of the actual counts); thus, select Proportion of Successes on the reference distribution app.

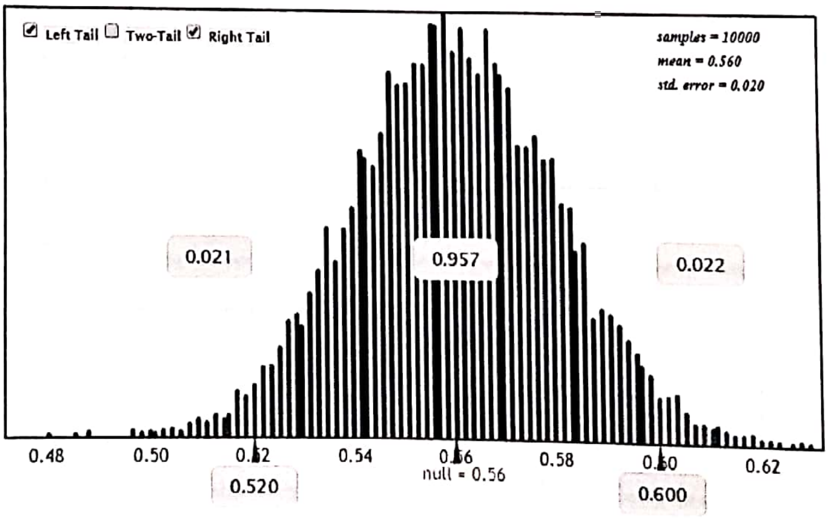
Guess #1

What proportion of the distribution sits between 51% and above 61%? 0.990
Using 51% and 61% as endpoints for the confidence interval is a little *too conservative*



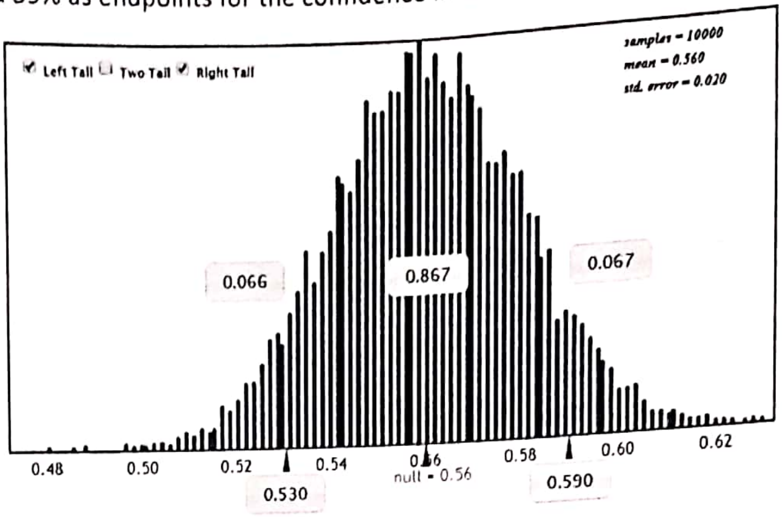
Guess #2

What proportion of the distribution sits between 52% and above 60%? 0.957
Using 52% and 60% as endpoints for the confidence interval *appears to be about right*



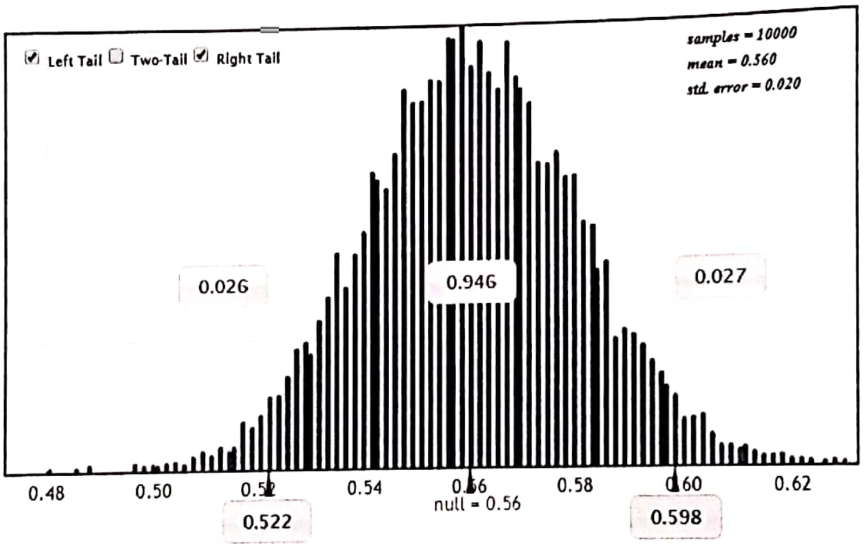
Guess #3

What proportion of the distribution sits below 53% and above 59%? 0.867
Using 53% and 59% as endpoints for the confidence interval appears to be *too liberal*



Refining Guess #2

What proportion of the distribution sits below 52.2% and above 59.8%? 0.9460
Using 52.2% and 59.8% as endpoints for the confidence interval would be even closer to 95%



$$95\%\ Confidence\ Interval:\ 52.2\% < \pi < 59.8\%$$

Margin-of-Error (MOE): $\pm$ 3.8%

$$Positive\ MOE\ = 59.8\% - 56\% = +3.8\%$$

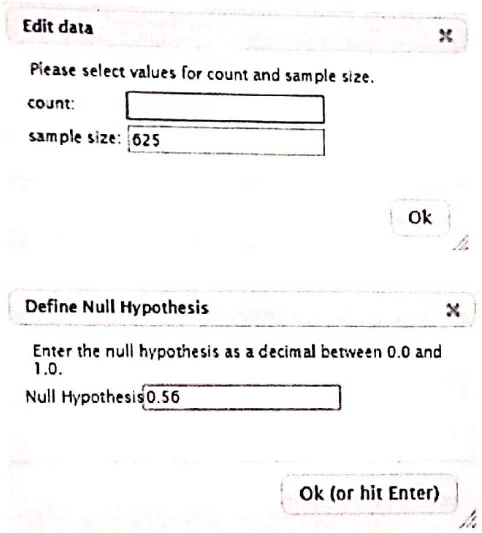
$$Negative\ MOE\ = 56\% - 52.2\% = -3.8\%$$

Approach #2: Using the Normal Theory Approach

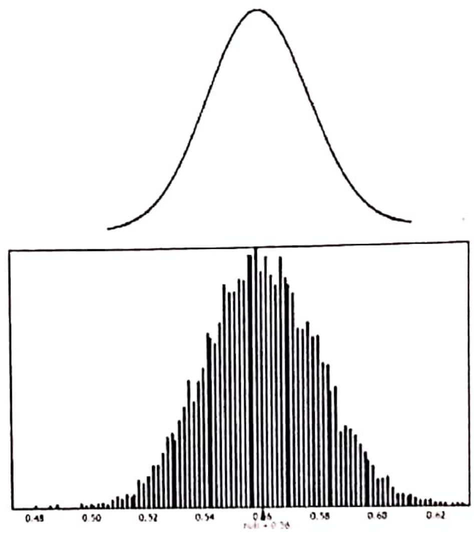
The limitations regarding the simulation model discussed in the previous section (regarding p-values) are present here as well, e.g. difference reference distributions will produce slightly different confidence intervals. The most common approach to computing a 95% confidence interval for a single proportion uses the **normal curve** (or bell-curve) approximation to the reference distribution.

The normal curve approximation works well for most situation. If the reference distribution is pushed too low (or too high) against the boundaries, the normal curve approximation should not be used. In this case, the confidence interval should be obtained directly from the binomial probability distribution.

Setup for Reference Distribution



Normal Curve Approximation



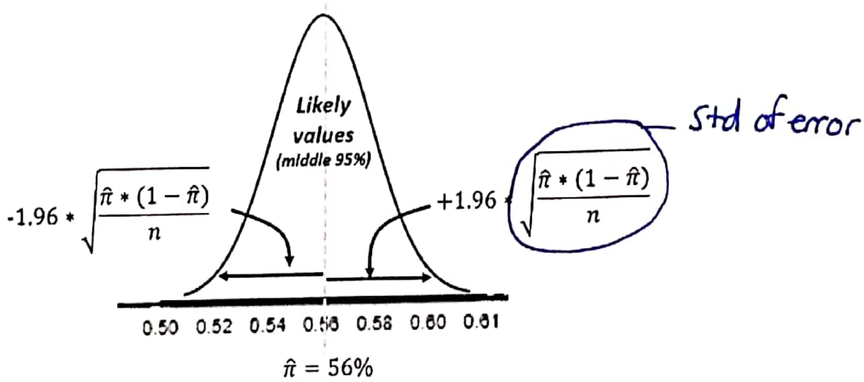
The center of the normal curve is placed at the point estimate, i.e. $\hat{\pi} = 56\%$, the proportion from the study. The sample size, i.e. $n = 625$, is also needed when using the normal curve approximation. The normal curve is approximating the binomial probability distribution which is based on an infinite number of repeated simulation.

Computing the 95% confidence interval via the normal theory approach

Quantity	Value	Description
Statistic (or point estimate)	$\hat{\pi}$	The proportion of interest from the study, i.e. $\hat{\pi} = 56\%$
Upper Margin-of-Error	$+1.96 * \sqrt{\frac{\hat{\pi} * (1-\hat{\pi})}{n}}$	The *positive* amount of acceptable inherent variation over repeated sampling for the statistic
Lower Margin-of-Error	$-1.96 * \sqrt{\frac{\hat{\pi} * (1-\hat{\pi})}{n}}$	The *negative* amount of acceptable inherent variation over repeated sampling for the statistic
Upper Endpoint	$\hat{\pi} + 1.96 * \sqrt{\frac{\hat{\pi} * (1-\hat{\pi})}{n}}$	The point on the high side of the distribution that separates the *likely* from *unlikely* values
Lower Endpoint	$\hat{\pi} - 1.96 * \sqrt{\frac{\hat{\pi} * (1-\hat{\pi})}{n}}$	The point on the low side of the distribution that separates the *likely* from *unlikely* values

Comments

- The quantity $\sqrt{\frac{\hat{\pi}*(1-\hat{\pi})}{n}}$ is commonly referred to as the **standard error**
- The value of 1.96 is used in these calculations so that the middle 95% is captured. Using 1.96 ensure the 5% rule is being used to separate likely from unlikely values



Margin-of-Error: $\pm 1.96 * \sqrt{\frac{\hat{\pi}*(1-\hat{\pi})}{n}} = \pm 1.96 * \sqrt{\frac{0.56*(1-0.56)}{625}} = \pm 0.0389 = \pm 3.9\%$

Upper Endpoint: 56% + 3.9% = 59.9%

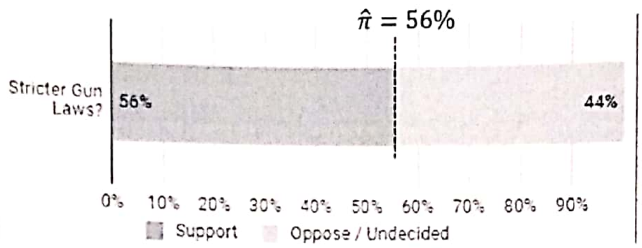
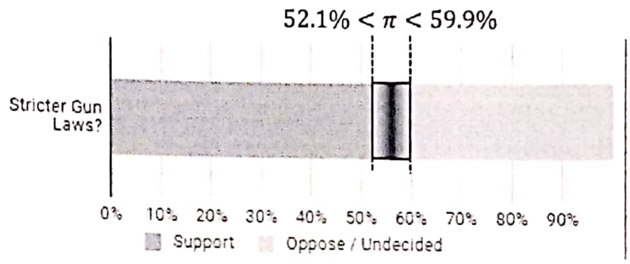
Lower Endpoint: 56% - 3.9% = 52.1%

95% Confidence Interval: $52.1\% < \pi < 59.9\%$

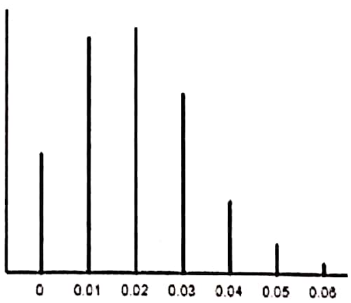
Interpretation: The proportion of *all* Minnesota Registered Voters that would support stricter gun laws in the US is likely to be between 52.1% up to 59.9%.

Comments

- The confidence interval identifies a range of likely values for the parameter which relates directly to the scope-of-inference -- not just the study outcome.

Outcome for 625 study participants	
Outcomes for all MN Registered Voters, i.e. the scope-of-inference	

- The normal approximation works well for most situation. The normal distribution fails when the reference distribution is pushed too close to smallest possible value, i.e. 0 (or too close to the largest possible value, i.e. n). The following is an example where the lower end of the normal approximation is truncated.

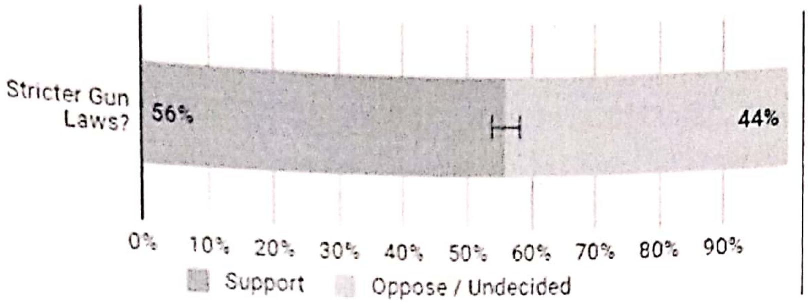


A commonly referred to guideline when using the normal approximation to these reference distributions is that the number of successes (and failures) should be larger than 10. This rule of thumb is usually stated as follows.

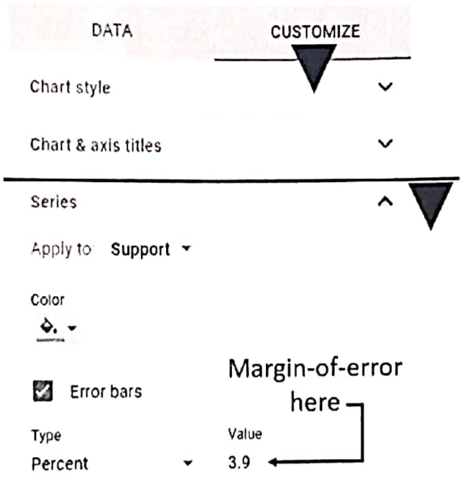
$$n * \hat{\pi} > 10$$

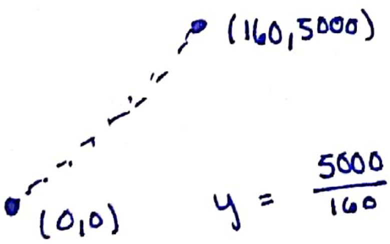
$$n * (1 - \hat{\pi}) > 10$$

- Most software packages make it easy to add the margin-of-error to a graph. The following was done in Google Sheets.



To add the margin-of-error, double-click on the segment of the graph for which the margin-of-error is to be added. Under Customize > Series menu, enter the margin-of-error in the box provided. The margin-of-error bars should be added to plot.

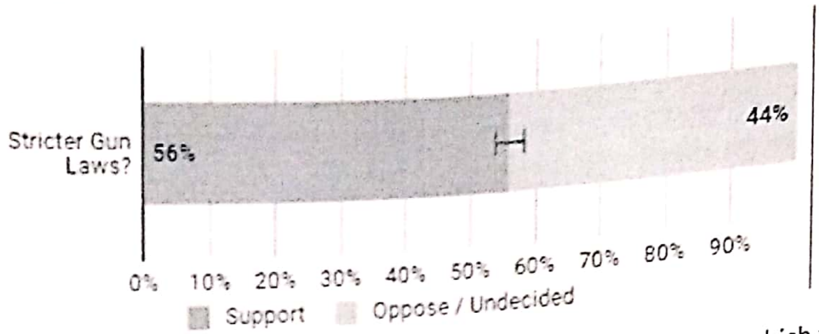




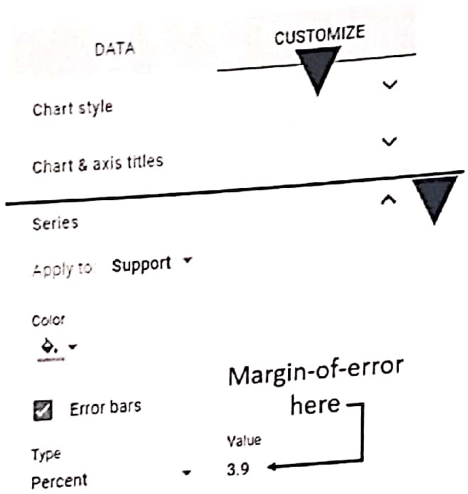
$y = \frac{5000}{160} x + 0$

$$n * (1 - \hat{\pi}) > 10$$

- Most software packages make it easy to add the margin-of-error to a graph. The following was done in Google Sheets.



To add the margin-of-error, double-click on the segment of the graph for which the margin-of-error is to be added. Under Customize > Series menu, enter the margin-of-error in the box provided. The margin-of-error bars should be added to plot.



(160, 5000)

(0, 0)

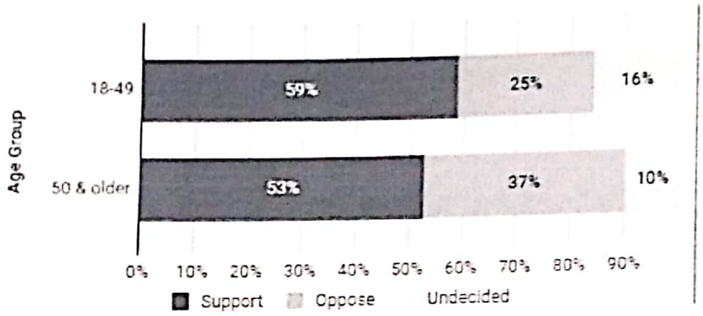
$$y = \frac{5000}{160} x + 0$$

Example 2.4.2 Consider the following breakdown from the data considered in the previous example. About 59% of the 299 people in the 18-49 age group supported stricter gun laws in the US, this dropped to 53% for the 324 people in the 50 & older age group.

Survey Question: Do you support or oppose stricter gun laws in the United States?

Age Group	Survey Response			Total
	Support	Oppose	Undecided	
18-49	176 (59%)	75 (25%)	48 (16%)	299
50 & older	172 (53%)	120 (37%)	32 (10%)	324
2 people refused to provide their age and have been excluded				

The data from the table above has been put into the following graphic using Google Sheets.



The following show the calculations for the margin-of-error for the Support survey response for both Age groups. The calculations done here were done in a spreadsheet.

$$\text{Margin-of-Error} = \pm 1.96 * \sqrt{\frac{\hat{\pi} * (1-\hat{\pi})}{n}}$$

Showing the calculations in a spreadsheet.

	Margin-of-Error
18-49:	= 1.96 * sqrt((0.59 * (1 - 0.59)) / 299)
50 & older:	= 1.96 * sqrt((0.53 * (1 - 0.53)) / 324)

	Margin-of-Error
18-49:	0.0557491707
50 & older:	0.05434635609

Margin of Error

- $18-49$ Age Group: $\pm$ 5.6%
- 50 & older Age Group: $\pm$ 5.4%

95% Confidence Interval

- 18 – 49 Age Group: $53.4\% < \pi_{18-49} < 64.6\%$
- 50 & older Age Group: $47.6\% < \pi_{50\,\&\,older} < 58.4\%$

Interpretation

- 18 – 49 Age Group: The proportion of *all* Minnesota Registered Voters whose age is between 18 and 49 that would support stricter gun laws in the US is likely to be between 53.4% up to 64.6%.
- 50 & older Age Group: The proportion of *all* Minnesota Registered Voters whose age is 50 & older that would support stricter gun laws in the US is likely to be between 47.6% up to 58.4%.

Comments

$$MOE = \pm 1.96 * \sqrt{\frac{\hat{\pi} * (1 - \hat{\pi})}{n}}$$

- In order to cut the margin-of-error in half, the sample size must be 4 times larger. This is because of the square root in the margin-of-error formula. The following example show the calculations for the 18-49 age group with 299 people and 299 x 4 = 1196 people.

Sample Size	Age Group	Margin-of-Error
299	18-49:	= 1.96 * sqrt((0.59 * (1 - 0.59)) / 299)
1196 (4 x larger)	18-49:	= 1.96 * sqrt((0.59 * (1 - 0.59)) / 1196)

Sample Size	Age Group	Margin-of-Error
299	18-49:	0.0557491707
1196 (4 x larger)	18-49:	0.02787458535

- With a little algebra, the formula for the margin-of-error can be expressed as follows. This expression can be used to determine how many people should be included in a study in order to achieve a certain margin-of-error.

$$n = \frac{\hat{\pi} * (1 - \hat{\pi})}{\left(MOE/_{1.96}\right)^2}$$

For example, suppose a $\pm$ 3% margin-of-error is desired and it is believe that the value of $\hat{\pi}$ is about 28%. The number of people required in this study to achieve a $\pm$ 3% margin-of-error is about 860.

$$n = \frac{\hat{\pi} * (1-\hat{\pi})}{\left(MOE/_{1.96}\right)^2} = \frac{0.28 * (1-0.28)}{\left(0.03/_{1.96}\right)^2} = 860.52$$