

# 7 Confidence Intervals

February 18, 2021 6:31 PM

FOM 11

Ch6: Statistics

## Day 7: Confidence Intervals

The Angus Reid Survey "Kids and Covid-19" of March 11, 2020 survey found:

- 54% of students said "Seeing Friends" was what they missed the most during lockdown.
- 27% said their biggest worry was "Missing the current school year".



These results carry a margin of error of  $\pm 4$  percentage points, 19 times out of 20

What does this mean?!

**Big Ideas:** confidence interval, margin of error, and confidence level can be used to interpret statistics based a *sample (subset)* in comparison to the *whole population of interest* (which the sample is supposed to represent).

Surveys are based on a subset of a population of interest ("target population"). The number of people included in the survey is the sample size.

- The margin of error is how many percentage points your survey result will differ from the *true value for the population* (i.e., if you had surveyed ALL of your target population).
  - It is usually expressed as a plus or minus percent, such as  $\pm 4$  for this survey.
- A confidence interval represents the accuracy of the sample's results. It is the interval in which the *true value (for the whole population of interest) is estimated to lie*, with a stated degree of error. The confidence interval is expressed using  $\pm$  notation, such as  $82\% \pm 4\%$ , or with ranges, like from 78 % to 86 %.
  - Confidence interval for "Miss Seeing Friends":  $54 \pm 4 = 50 \text{ to } 58$
  - Confidence interval for "Worry about missing current school year":  $27 \pm 4 = 23 \text{ to } 31$
- The Confidence Level is the likelihood that the result for the "true" population lies within the range of the confidence interval.

$$19 \text{ times out of } 20$$
$$\frac{19}{20} = 95\%$$

margin of error:  $\pm 4\%$ 

**Example 1a:** Referring to the Angus Reid study results for the question "Now thinking about life outside your schoolwork, what are you doing to fill your time?", what percentage of students aged 16-17 replied "Reading/Drawing/Music/Hobbies"? 53%



What is the **confidence interval** for this (i.e., if you were to project this result onto the whole population)?  $53 \pm 4\% = 49\% \text{ to } 57\%$

**Example 1b:** Regarding the question "Would you say that things with your friends are better, the same, or worse than they were since you've had to stay home?", what is the confidence interval for students aged 16-17 who replied "better"?

6%  $6\% \pm 4\% = 2\% \text{ to } 10\%$

How many students said "better"? (You need to multiply by the sample size!)

Sample size = 194 students

$$2\% \text{ of } 194 = \frac{2}{100} \times 194 = 3.9 \approx 4 \text{ students}$$

$$10\% \text{ of } 194 = \frac{10}{100} \times 194 = 19 \text{ students}$$

**Example 2:** A survey of 320 users of the skateboard park indicates that 40% of them would like the parks board to extend the evening use of the facility. This survey is considered accurate to within 5.4%, 18 times out of 20.

a) What is the margin of error?  $\pm 5.4\%$

b) What is the **confidence interval**?  
 $40\% \pm 5.4\% = 34.6\% \text{ to } 45.4\%$

c) What is the confidence level?  
 $18/20 = 90\%$

d) Calculate the range of people who want to extend the evening hours: represented in the survey

$$34.6\% \text{ of } 320 = 34.6 \div 100 \times 320 = 110.7 \approx 111 \text{ people}$$

$$45.4\% \text{ of } 320 = 45.4 \div 100 \times 320 = 145.28 \approx 145$$



**Example 3:** Polling organization in Canada frequently survey samples of the population to gauge voter preference prior to elections. People are asked:

- "If an election were held today, which party would you vote for?"
- If they say they don't know, then they are asked "Which party are you leaning toward voting for?"



The results of 3 different polls taken during the first week of Nov. 2010 are shown (Stephen Harper eventually won the election).

Polling Organization & Data	Conservative (%)	Liberal (%)	NDP (%)	Bloc Quebecois (%)	Green Party (%)	Undecided (%)
Ekos	29	29	19	9	11	12.6
sample size, 1815 margin of error, $\pm 2.3\%$						
Nanos	37	32	15	11	5	19.2
sample size, 844 margin of error, $\pm 3.4\%$						
Ipsos	35	29	12	11	12	n.a.
sample size, 1000 margin of error, $\pm 3.1\%$						

source: <http://www.sfu.ca/~aheard/elections/polls.html>

a) How does the sample size used in the poll affect the margin of error in the reported results?

*the larger the sample size,  
the less the margin of error,  
(more accurate!),*

b) Compare the confidence intervals for the Conservative party for each of the polls. How does the sample size used in the poll affect the confidence interval?

*the smaller the intervals!*

**Example 4:** In a factory that produces baseballs, a quality control engineer takes a random sample of baseballs daily and measures their mass to determine their mean mass. If the mean mass of the random sample is 144.7g to 145.3g, then the production equipment is running correctly. If the mean mass of the sample is outside the acceptable level, the production equipment is shut down and adjusted. For the quality control tests, identify:



a) the confidence interval

$$144.7 \text{ g to } 145.3 \text{ g}$$

$$\begin{array}{r} 145.3 \\ - 0.3 \\ \hline 145 \end{array}$$

$$\rightarrow 145 \pm 0.3 \text{ g}$$

b) margin of error

$$\begin{array}{r} 145.3 \\ - 144.7 \\ \hline 0.6 \end{array}$$

subtract then divide by 2

$$\Rightarrow 0.6 \div 2 = 0.3 \Rightarrow \pm 0.3$$

c) The quality control engineer used this chart when conducting random sampling:

Confidence Level	Sample Size Needed
99%	110
95%	65
90%	45

What is the relationship between confidence level and sample size?

the larger the sample size,  
the higher the confidence level!

**Assignment:** Confidence Interval Worksheet