5.6

Confidence Intervals [5.6]

In Summary

Key Ideas

- It is often impractical, if not impossible, to obtain data for a complete population. Instead, random samples of the population are taken, and the mean and standard deviation of the data are determined. This information is then used to make predictions about the population.
- When data approximates a normal distribution, a confidence interval indicates the range in which the mean of any sample of data of a given size would be expected to lie, with a stated level of confidence. This confidence interval can then be used to estimate the range of the mean for the population.
- Sample size, confidence level, and population size determine the size of the confidence interval for a given confidence level.

Need to Know

- A confidence interval is expressed as the survey or poll result, plus or minus the margin of error.
- The margin of error increases as the confidence level increases (with a constant sample size). The sample size that is needed also increases as the confidence level increases (with a constant margin of error).
- The sample size affects the margin of error. A larger sample results in a smaller margin of error, assuming that the same confidence level is required.

sample size — how many people/objects

margin of error

For example:

- A sample of 1000 is considered to be accurate to within ±3.1%, 19 times out of 20. *— confidence level = 95%*
- A sample of 2000 is considered to be accurate to within ±2.2%, 19 times out of 20.
- A sample of 3000 is considered to be accurate to within ±1.8%, 19 times out of 20.

confidence interval comes from the result ± margin of error

ex   10g ± 2g , 90% of the time

Ans: With 90% confidence we can say that the packages range between 8g and 12g.

10g−2g   10g+2g

(±)

— to —

The Effect of Sample Size on Margin of Error and Size of Confidence Interval

The 2011 Canadian federal election took place on May 2nd, 2011. A large number of opinion polls were conducted in the days leading up to the election.

(link: http://en.wikipedia.org/wiki/Opinion_polling_in_the_Canadian_federal_election,_2011)

The table below gives the results of four of these polls. The data represents the percent of the sample who would cast their vote for each party. Some of the results do not add up to 100% because votes for minor parties and independents are not included in the table. The results are accurate to the stated margin of error 19 times out of 20.   — confidence level = 95%

(±)

Polling Firm	Date of Poll	Number in Sample	Con	Lib	NDP	BQ	Green	Margin of Error
Harris Decima	Apr 27	1011	35%	22%	30%	5%	7%	±3.1%
Angus Reid	Apr 29	2197	37%	19%	33%	6%	4%	±2.2%
Nanos Research	Apr 30	1048	37.0%	22.7%	30.6%	5.5%	3.2%	±3.0%
EKOS Research	May 1	2690	33.9%	21.0%	31.2%	6.4%	6.0%	±1.8%

a) Look at the columns for **Number in Sample** and **Margin of Error**. How does the sample size affect the margin of error?

smaller the sample size, the bigger the margin of error

b) For each polling firm, calculate the 95% confidence interval for the percent of Canadian voters who would vote for the New Democratic Party (NDP) and complete the table. How does sample size affect the range of values in the confidence interval?

— to —

NDP

Polling Firm	95% Confidence Interval
Harris Decima	30% ±3.1% = 26.9% to 33.1%
Angus Reid	33% ± 2.2% = 31.8% to 35.2%
Nanos Research	30.6% ±3% = 27.6% to 33.6%
EKOS Research	31.2% ± 1.8% = 29.4% to 33%

← range will happen 19 times out of 20, or 95% of the time

c) For each polling firm, calculate the 95% confidence interval for the percent of Canadian voters who would vote for the Conservative Party and complete the table below.

copy from prev page

Polling Firm	Confidence Interval <u>Conservative</u>	Confidence Interval NDP
Harris Decima	$35 \pm 3.1 = 31.9\%$ to 38.9%	26.9% to 33.1% *overlaps*
Angus Reid	$37 \pm 2.2 = 34.8\%$ to 39.2%	31.8% to 35.2% *"*
Nanos Research	$37 \pm 3 = 34\%$ to 40%	27.6% to 33.6% *(no overlap)*
EKOS Research	$33.9 \pm 1.8 = 32.1\%$ to 35.7%	29.4% to 33% *overlap*

i) In which of the polling firms is there no overlap of percents in the confidence intervals?   *Nanos Research*

In this case, since there is no overlap, we can say with 95% confidence that the percent of Canadian voters who will vote Conservative is greater than the percent of Canadian voters who will vote NDP.

ii) In which of the polling firms is there an overlap of percents in the confidence intervals?   *all the rest*

In these cases, since there is an overlap, we <u>cannot</u> say with 95% confidence that the percent of Canadian voters who will vote Conservative is greater than the percent of Canadian voters who will vote NDP.

d) In which of the polls can we say with 95% confidence that the NDP will receive a higher percent of the vote than the Liberal Party?

e) In which of the polls can we say with 95% confidence that the Bloc Quebecois will receive a higher percent of the vote than the Green Party?

Example 2:

In order to determine the mean $\overset{\mu}{\text{mass}}$ of a type of chocolate bar produced at a factory, sampling is done and the following statement is made by the company.

"The mean mass of chocolate bars produced at our factory is $\overset{\mu}{100.4}$ grams ± 0.4 grams. — margin of error
The results are accurate 18 times out of 20." — 90% confidence level

a) Determine the confidence level.

18 times out of 20 = 90%

b) Determine the confidence interval.

$100.4g \pm 0.4g = \overline{100g}$ to $\overline{100.8g}$

c) State the margin of error.

$\pm 0.4g$

d) Is it likely that the mean mass of chocolate bars produced at this factory is 99.5 grams? Explain.

not likely since we are 90% confident that chocolate bars are between 100g to 100.8g (99.5g is not in the confidence interval)

e) Is it possible that the mean mass of chocolate bars produced at this factory is 99.5 grams? Explain.

yes it is possible since 10% of the time the chocolate bars ~~are~~ could be outside the confidence interval

Practice pg 274 # 4, 6, 7, 8, 9, 10
                                     ∟ use answer key to help.

(10b)

$\frac{18}{20}$ = 90% conf. level   → $\frac{19}{20}$ = 95% conf. level (only 5% of time will fall outside range)

10% of time will fall outside 30-34%   30% to 34% vote Conservative   29% to 35% vote Conservative

32% ± 2%   32% ± 3%
   ↑   ↑
margin of error   larger margin of error

0-29, 35-100%