5.6 Confidence Intervals (Concepts 29 and 30)

Suppose: A telephone survey of 600 randomly selected people was conducted in an urban area. The survey determined that 76% of people, from 18 to 34 years of age, have a social networking account. The results are accurate within plus or minus 4 percent points, 19 times out of 20.

This means that if 600 people between the ages of 18 and 34 years of age were surveyed 20 times, 19 of those times (or 95% of the time), somewhere between 76% ± 4% (72% to 80%) would say they have a social networking account.

Margin of error – The possible difference between the estimated result from a survey and the true value of the population

Ex: ± 4% (from example above)

Confidence Interval – The interval in which the true value is estimated to fall

Ex: 76% ± 4 or 72 to 80

Confidence Level – The likelihood that the results for the entire population will fall within the predicted confidence interval

"If the survey were conducted 100 times then 95 times out of 100 the percent of people in the population with a social networking account would be from 72%–80%"

Ex: the results are accurate 19 times out of 20, which is 95% of the time. (from example above)

The confidence level is 95%. **(expressed as a** percentage **)**

Example 1: Suppose that if the total population of 18 to 34 year olds in this same suburban area where the telephone survey was conducted is 92 500, how can the above results be interpreted to indicate the **true** number of people who would have a social networking account?

- Use the **confidence interval** of 76% ± 4% to find the range of 18 to 34 year olds who would likely have an account.

72% of 92500
$= 0.72 \times 92500$
$= 66\,600$

80% of 92500
$= 0.8 \times 92500$
$= 74000$

Therefore it can be said, within 95% confidence, that 66 600 to 74000 people, in a population of 92 500 people from ages 18 to 34, have a social networking account.

Example 2- Analyzing the effect of SAMPLE SIZE on margin of error and confidence intervals

Polling organizations 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 three different polls taken during the first week of November, 2010, are shown below. The results of each poll are considered accurate 19 times out of 20.

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, ±2.3%					
Nanos	37	32	15	11	5	19.2
sample size, 844	margin of error, ±3.4%					
Ipsos	35	29	12	11	12	n.a.
sample size, 1000	margin of error, ±3.1%					

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

What effect does the SAMPLE SIZE (n) used in a poll have on:

the **MARGIN OF ERROR** in the reported results?

- as the sample size increases, the margin of error decreases

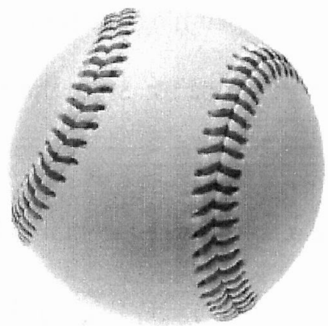
- A larger sample *should be* a better indicator of how the population might vote, therefore resulting in a smaller margin of error.

the **CONFIDENCE INTERVAL** in the reported results?

- If the margin of error decreases as the sample size increases, what do you expect to happen to the confidence interval as the margin of error decreases?

If polls are conducted using the same confidence level, as the sample size increases, the **range** in the confidence interval decreases

Example 3 Analyzing the effect of confidence levels on sample size



To meet regulation standards, baseballs must have a mass from 142.0 g to 149.0 g. A manufacturing company has set its production equipment to create baseballs that have a mean mass of 145.0 g.

To ensure that the production equipment continues to operate as expected, the quality control engineer takes a random sample of baseballs each day and measures their mass to determine the mean mass. If the mean mass of the random sample is 144.7 g to 145.3 g, 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. The quality control engineer refers to the table below when conducting random sampling.

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

a) What is the confidence interval and margin of error the engineer is using for quality control tests?

142-149 g → confidence interval $\bar{x} = 145$ g 145 ± 0.3 → margin of error.

b) Interpret the table.

- In order to be confident that, 99 out of 100 times, the mean mass of the sample measures from 144.7 g to 145.3 g, the engineer needs to take a random sample of 110 baseballs from the production line.
- In order to be confident that, 95 out of 100 times, the mean mass of the sample measures from 144.7 g to 145.3 g, the engineer needs to take a random sample of 65 baseballs from the production line.
- In order to be confident that, 90 out of 100 times, the mean mass of the sample measures from 144.7 g to 145.3 g, the engineer needs to take a random sample of 55 baseballs from the production line.

What is the relationship between CONFIDENCE LEVEL and SAMPLE SIZE?

For a constant margin of error, as the confidence level increases, the size of the sample needed to attain that level of confidence ___increases___. To have a greater confidence that the baseballs meet quality standards, the engineer must use a ___larger___ sample.

Example 4)

a. A poll was conducted about an upcoming election. The results are considered accurate within ±4 percent points, 9 times out of 10. State the confidence level.

90% confidence level

$9/10 = 90\%$

b. The results of a survey have a confidence interval of 29% to 37%, 9 times out of 10. Determine the margin of error.

$\frac{29+37}{2} = 33 \pm 4\%$

Margin of error $= \pm 4\%$

c. A poll was conducted about an upcoming election. The result that 52% of people intend to vote for one of the candidates is considered accurate within ±8.0 percent points, 19 times out of 20. State the confidence interval.

46-60% = Confidence interval

$52 \pm 8\%$

$52+8=60$

$52-8=46$

d. In a recent survey of high school students, 65% of those surveyed said they would vote for Chuck as student council VP. The survey is considered accurate to within ±8.2 percent points, 19 times out of 20. If a high school has 2000 students, state the range of the number of votes Chuck should expect.

Confidence interval $65\% \pm 8.2\%$

$56.8\% - 73.2\%$

$0.568 \times 2000 = 1136$

$0.732 \times 2000 = 1464$

Chuck can expect between 1136 - 1464 students to vote

e. Match the correct sample size with the appropriate margin of error.

Sample Size	Margin of Error
40	0.09
100	0.29
200	0.14

↓ sample size margin of error ↑

5.6 Assignment: page 274 #1,3,4,6,9,10