Key

Sampling and Error

Vocabulary

1. Bias: A sample of size *n* is random (or unbiased) when every possible sample of size *n* had an equal chance of being selected. If a sample is biased, then information obtained from it may not be reliable.

   - Example: To find out how people feel about mass transit, people at a commuter train station are asked their opinion. Does this situation represent a random sample?

     No, the sample includes only people who actually use a mass-transit facility. The sample does not include people who ride bikes, drive cars, or walk.

Practice: Determine whether each situation would produce a random sample. Write *yes* or *no* and explain your answer.

1. Asking people in Phoenix, Arizona, about rainfall to determine the average rainfall for the United States.

   No, surveying one location would not be representative of the entire U.S.

2. Obtaining the names of tree types in North America by surveying all the U.S. National Forests.

   No, may not be a bad sample for U.S. but this does not take into consideration Mexico & Canada.

3. Surveying every tenth person who enters the mall to find out about music preferences in that part of the country.

   Yes, even though you are not getting non-mall shoppers no reason to think this effects music preference

4. Interviewing country club members to determine the average number of televisions per household in the community.

   No, country club members would be wealthier than norm.

5. Surveying all students whose ID numbers end in 4 about their grades and career counseling needs.

   Yes, random

6. Surveying parents at a day care facility about their preferences for brands of baby food for a marketing campaign.

   No, excluding stay at home moms and kids at in home daycares

7. Asking people in a library about the number of magazines to which they subscribe in order to describe the reading habits of a town.

   No, library will have people who read more

Unit 4 1.9

Vocabulary

Margin of Error: The **margin of sampling error** gives a limit on the difference between how a sample responds and how the total population would respond.

Margin of Error: an amount (usually small) that is allowed for in case of miscalculation or change of circumstances.

- "p" : the percent of people in a sample responding in a certain way
- "n" : the size of the sample
- Margin of error formula (with a 95% confidence interval): $ME = 2\sqrt{\frac{p(1-p)}{n}}$
  - In future courses you will use a variety of confidence levels, for our purposes we will always use a 95% level of confidence.
  - To change levels of confidence you would change the "2" in the formula. (95% is within 2 standard deviations of the mean, see 1.7 notes)

Example 1: In a survey of 4500 randomly selected voters, 62% favored candidate A. What is the margin of error?

$ME = 2\sqrt{\frac{p(1-p)}{n}}$ $\quad ME = 2\sqrt{\frac{0.62*0.38}{4500}}$ $\quad ME = 0.01447 \text{ or about } 1.4\%$

The margin of error is about 1.4%. This means that there is a 95% chance that the percent of voters favoring candidate A is $62\% \pm 1.4\% \text{ or } 60.6\% < p < 63.4\%$.

Example 2: In a survey of 1068 Americans, 673 stated that they had answering machines. Find the a) percent of people that stated they have an answering machine, b) the margin of error to the nearest tenth of a percent, and c) the confidence interval.

a) $p = \frac{673}{1068}$ = 63% $\quad$ b) $ME = 2\sqrt{\frac{0.63*0.37}{1068}}$ = 0.0295 or 3% $\quad$ c) $63\% \pm 3\%$

Calculating the margin of error

1. Determine the percentage of responses.
   - $p = \frac{x}{n}$, where p is the proportion of x successes in a sample size of n.
2. Fill in formula and calculate the percentage.
3. If they just ask for the margin of error, write that percentage. If they ask for the confidence interval, write your answer as $P\% \pm ME\%$

In problems 1-4, find the margin of sampling error to the nearest tenth of a percent.

1. p=45%, n=350

$ME = 2\sqrt{\frac{(.45)(.55)}{350}}$

ME = 5.32%

2. p=12%, n=1500

$ME = 2\sqrt{\frac{(.12)(.88)}{1500}}$

ME = 1.68%

3. p=86%, n=600

$ME = 2\sqrt{\frac{(.86)(.14)}{600}}$

ME = 2.83%

4. A study of 50,000 drivers in Indiana, Illinois, and Ohio showed that 68% preferred a speed limit of 75 mph over 65 mph on highways and country roads. What is the 95% confidence interval?

$$ME = 2\sqrt{\frac{(.68)(.32)}{50{,}000}}$$

$$68\% \pm .42\%$$

$$ME = .00417 \rightarrow .42\%$$

5. The Hartford Insurance Company wants to estimate the percentage of drivers who text or talk on their phone while driving. A random sample of 850 drivers found that 544 who text or talk on the phone while driving. Find the 95% confidence interval for the percentage of all drivers who text or talk on the phone while driving.

$$p = \frac{544}{850} = 64\%$$

$$ME = 2\sqrt{\frac{(.64)(.36)}{850}}$$

$$ME = .03292\ldots \Rightarrow 3.29\%$$

$$C.I. \Rightarrow 64\% \pm 3.29\%$$

6. When 500 college students are randomly selected and surveyed, if was found that 235 of them kept a car at college. Calculate the 95% confidence interval for the percentage of all college students that have a car at college.

$$p = \frac{235}{500} = 47\%$$

$$ME = 2\sqrt{\frac{(.47)(.53)}{500}}$$

$$= 4.46\%$$

$$47\% \pm 4.46\%$$

7. The Spaulding Corporation wants to estimate the proportion of golfers that are left-handed. Out of 80 golfers that were randomly surveyed they found that 9 were left-handed. Calculate the 95% confidence interval for the proportion of all golfers that are left-handed.

$$p = \frac{9}{80} = 11.25\%$$

$$ME = 2\sqrt{\frac{(.1125)(.8875)}{80}}$$

$$= 7.07\%$$

$$11.25\% \pm 7.07\%$$

8. A testing company considers a multiple-choice question easy if 80% of responses are correct. A random sample of 250 responses to one particular question includes 79% correct responses. Construct the 95% confidence interval for the true percentage of correct responses. Is it likely that this question is really easy?

$p = 79\%$ $\quad ME = 2\sqrt{\frac{(.79)(.21)}{250}} = 5.15\%$

$79\% \pm 5.15\%$

9. The Ford Motor Company took a random sample of 1220 households in which 1054 owned a vehicle. Construct the 95% confidence interval for the percentage of all households that own a car.

$p = \frac{1054}{1220} = 86.4\%$ $\quad ME = 2\sqrt{\frac{(.864)(.136)}{1220}}$

$= 1.96\%$

$86.4\% \pm 1.96\%$

10. In a study of store checkout scanners, 1234 items were checked and 20 of them were found to be overcharges. Using the sample data construct the 95% confidence interval for the proportion of all such scanned items that are overcharges.

$p = \frac{20}{1234} = 1.6\%$ $\quad ME = 2\sqrt{\frac{(.016)(.984)}{1234}}$

$ME = .71\%$

$1.6\% \pm 0.71\%$