**Margin of Error**

**Part I Do you Tweet?**

The Pew Internet Research Project, in an internet article titled *The Demographics of Social Media Users – 2012*, reported the results from a landline and cellphone survey of internet users. The table below is taken from the online article.

|  |  |
| --- | --- |
| Retrieved from <http://www.pewinternet.org/2013/02/14/the-demographics-of-social-media-users-2012/>. | Using the table at left, answer the following questions:1. What was the population of interest?
2. How many people were in the sample?

 **\_\_\_\_\_\_\_\_\_\_\_**1. What percent of all internet users use Twitter?

 **\_\_\_\_\_\_\_\_\_\_\_** 1. What **margin of error** is reported?

 **\_\_\_\_\_\_\_\_\_\_\_**1. What do you think this margin of error means?
 |

**Part II Paper Bag Population - Exploring the Margin of Error**

Your teacher has a paper bag that contains an entire population of colored beads. What colors are the beads in the bag? How many beads are in the bag? Are there any blue beads in the bag? If so, what proportion of the beads is blue? How do we answer these questions if we cannot see the entire population?

In real life, we are usually unable to make observations about an entire population. It is either not feasible or simply not practical to try to gather data from the whole population. To learn about the population, we take one or more samples from the population. If the sample(s) is representative of the population, we can make inferences about the population.

The paper bag population simulates a real-life population. We cannot see the population, but we may take samples from the population.

* What type of sample of our paper bag population should we take to ensure that our sample is representative?

Let’s conduct a statistical investigation.

 **Formulate a question**

What proportion of the beads in the bag is blue?

**Design and implement a plan to collect data**

We will take random samples of size 25 from the paper bag population. The sample proportion of blue beads will be our estimate of the proportion of blue beads in the paper bag population. We use the symbol $\hat{p}$ (p-hat) to represent our estimate of the population proportion.

 Steps:

* Shake the bag to mix the beads. (Why are we doing this?)
* One student at a time should reach into the bag and remove 25 beads. If too many are selected, randomly pick beads from the sample to return to the population bag. If too few are selected, randomly pick the needed amount from the population bag.
* Count the number of blue beads in the sample and determine the proportion of blue beads in the sample.
* Plot this proportion on the class dot plot.
* Repeat until each student in the class has taken at least one turn.

What is your sample proportion of blue beads? $\hat{p} $(p-hat) = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

This value will be your estimate of the proportion of blue beads in the paper bag population.

**Analyze the data by measures and graphs**

Copy the class dot plot here:

**Simulation of the Proportion of Blue Beads in the PaperBag Population**

**Sampling Distribution with Samples of Size 25**



 Proportion of Blue Beads in Sample

With your class, calculate the mean and standard deviation of the sampling distribution.

 Mean = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Standard Deviation = \_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Interpret the results in the context of the original question**

We know that samples vary. In the sampling distribution above, we can see that our sample proportions vary in a predictable way. In our work with sampling distributions we found that:

* The shape of the sampling distribution is: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
* The center of the sampling distribution equals \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
* The spread of the sampling distribution is less when \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
* We expect \_\_\_\_\_\_\_% of our sample proportions to be within 2 standard deviations of the true population proportion.

|  |
| --- |
| **Margin of Error**One method of constructing an interval is to use a sample statistic ± a margin of error. The margin of error is 2 times the standard deviation of the sampling distribution.$$sample statistic\pm margin of error $$$$\overbar{x} \pm 2∙std Dev or \hat{p}\pm 2∙std Dev $$We use the reasoning that, if the sample statistic is likely to be within 2 standard deviations of the center of the sampling distribution, then the center of the distribution is also likely to be within 2 standard deviations of the sample statistic. Remember that the population parameter is approximately equal to the mean of the sampling distribution, so the population parameter is likely to be within two standard deviations of our sample statistic. In other words, we conclude that the population parameter is likely to be in this interval. (Or at least we expect 95% of the intervals created by this method to include the population parameter.)  |

* What is our margin of error?

Margin of error = $ 2∙std Dev $

 = $2∙$ \_\_\_\_\_

 = \_\_\_\_\_\_\_\_\_

* Report a margin of error using your sample proportion as the estimate for the population proportion. Your estimate should be given as:

$$\hat{p}\pm margin of error $$

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

* The true population proportion of blue beads in the paper bag population lies between what two values?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Part III Back to Tweeting**

In the Tweeting survey, a percent rather than a proportion is used.

1. What is the estimated value of the percent of internet users that use Twitter? \_\_\_\_\_\_\_
2. What is the reported margin of error? \_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. Write a sentence interpreting the margin of error for this survey.

**Part IV Sample Problems**

Problem 1 Hours of Homework

How many hours do American high school students spend per week on homework? To answer this question, a class of 28 students each took random samples of 25 United States high school students from the data base at Census at School (<http://www.amstat.org/censusatschool/index.cfm>). The dot plot below shows the distribution of the averages for their samples.



1. Based on the standard deviation of the sampling distribution above, what is the estimated margin of error?
2. One of the students in the class had a sample with a mean of 11.0 hours. Based on this sample mean, we anticipate the actual average number of hours spent on homework by United States high schools students to fall within what interval?
3. If each of the 28 students constructed his own interval by calculating $\overbar{x} \pm 2∙std Dev$ , would we expect the actual population mean to fall within each of the 28 intervals? Explain.
4. If the students had taken samples of size 50 instead of samples of size 25, how would this change the interval that you found in (b) above?

Problem 2 President Obama’s Job Approval Rating





From Gallup: <http://www.gallup.com/poll/113980/Gallup-Daily-Obama-Job-Approval.aspx>

1. What was the population of interest in this statistical study?
2. How was the sample selected?
3. How many people were in the sample?
4. Assuming that the sample of 1500 adults who were polled is representative of the American population, we infer that the actual percentage of Americans who approve of the job President Obama is doing as president is likely to be between what two values?
5. Is it likely that the actual percentage of Americans that approve of the job that President Obama is doing as president is 50% or greater? Explain.

1. If Gallop took a smaller sample, how would this change your answer to part (d)? Explain.

**Teacher Notes**

New Illinois Learning Standards addressed by this activity:

Content Standards

|  |  |  |  |
| --- | --- | --- | --- |
| Major | S | IC.4 | Make inferences and justify conclusions from sample surveys, experiments, and observational studies. Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.\* |

Primary Math Practices

MP 1 Make sense of problems and persevere in solving them.

MP 2 Reason abstractly and quantitatively.

MP 4 Model with mathematics.

MP 5 Use appropriate tools strategically.

MP 6 Attend to Precision

MP 7 Look for and make use of structure.

Acknowledgements

Duggan, M. & Brenner, J. (2013, February 14). *The Demographics of Social Media Users – 2012*. Retrieved from <http://www.pewinternet.org/2013/02/14/the-demographics-of-social-media-users-2012/>.

Census at School <http://www.amstat.org/censusatschool/>

Thank you to Roxy Peck for proof reading the two sample problems and providing guidance on the simulation to determine and explanation of the margin of error.

Guidance for Activity

The paper bag population may consist of any uniformly-shaped, small item. Beads, small Legos, or BBs all work well. As the instructor, it helps if you have a known proportion of one color in the bag. Do not let the proportion be too small unless you have a large population and can have a larger sample size. (Your sample size should be less than 10% of the population.) A larger population helps the students to realize that counting the entire population would not be a good method. A suggested population is to have 5,000 beads where 35% are white and 65% are blue.

You may want to assign students to read the internet article, *The Demographics of Social Media Users – 2012* as part of their work for S.IC.3 and 6.

An online applet can be used to illustrate margin of error. Explain that since we expect 95% of the sample proportion to be within 2 standard deviations of the population proportion (the center of the sampling distribution), then we expect the population proportion to be within 2 standard deviations of most of the sample proportions. We use the sample proportion +/- a margin of error to estimate the population proportion. When the sample proportion is used as a prediction of the population proportion, the symbol $\hat{p}$ (p-hat) is used. Students can either think of the margin of error as

 sample statistic ± margin of error

 or as

 $\hat{p} $± margin of error $\overbar{x}$± margin of error

where

 margin of error = 2$∙standard deviation of the sampling distribution ( std dev of estimate)$