**Summer Jobs**

A Margin of Error Activity

Mrs. Johnson, principal of Schueler High School, claims that at least 50% of the junior class has a summer job. Tomas, a junior at Schueler High School, would like to verify Mrs. Johnson’s claim. He takes a simple random sample of size 40 of students from the junior class. 18 out of the 40 students say they have a summer job. Does Tomas’ sample proportion provide evidence that the principal’s claim is incorrect? Using Tomas’ random sample, calculate a margin of error for the actual population proportion of juniors at Schueler High School that has a summer job.

* What is the population? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

* What is the population characteristic of interest? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

* What is the sample? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

* What is the sample statistic and, what is its value? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Let’s investigate this using a simulation. In a simulation, we assume that the principal’s claim is true. Suppose that 50% of juniors have a summer job. If we take random samples from the population of all juniors at Schueler High School, we expect that the proportion of juniors who have a summer job, while close to 0.5 (50%), will vary. Some samples would have a proportion that is more than 0.5, and some would have a proportion that is less than 0.5. Will 18/40 be one of the values that we are likely to get if at least 0.5 of juniors have a summer job?

* Describe how a simulation could be carried out using colored chips.
* Describe how a simulation could be carried out by tossing a coin.

**The Simulation**

In our simulation, we will use a TI-84 calculator to simulate tossing a coin 40 times. Below are the instructions for finding the sample proportion for a random sample of size 40 from a population with a population proportion of 0.5.

1. On your TI-84 Calculator:

First, reseed your calculator:

Enter a four digit number of your choice into your TI-84 then

 STO→

MATH → PRB

 1: rand

 ENTER.

1. Now, select a random sample of size 40 and find the sample proportion:

 Select MATH

* PRB

7: randBin(

Input: randBin(40, .5) / 40

randBin( ) will generate the number of times in 40 trial an outcome with a probability of 0.5 ( similar to tossing a coin) occurs. We then divide this number by 40 to get the proportion.

1. Record the proportion from your calculator’s simulated sample in the table below.
2. Repeat steps 2 and 3 three more times.

Your Simulated Samples

|  |  |
| --- | --- |
| Sample | Proportion |
| Sample 1 |  |
| Sample 2 |  |
| Sample 3 |  |
| Sample 4 |  |

* Based on your proportion in Sample 1, if you had tossed a coin

40 times, how many times would you have ended up with heads? \_\_\_\_\_\_\_\_\_

This represents the number of juniors in a random sample of 40 students who have a summer job.

1. Plot your four sample proportions on the class dot plot.
2. Record a copy of the class dot plot below.



**Analysis**

* Describe the shape, center, and spread of the distribution of the *Simulated Sample Proportions*.

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

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

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

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

* Approximately what proportion is at the center of the class dot plot? \_\_\_\_\_\_\_\_
* Are you surprised that this proportion is at the center? Why or why not?

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

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

* As a class, use a calculator to find the mean and standard deviation for the *Simulated Random Samples Distribution*.

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

* How many of the class samples were at 0.45 (18/40) or below? \_\_\_\_\_\_\_\_
* How many total samples were taken by your class? \_\_\_\_\_\_\_\_\_\_
* What proportion of your class samples is at 0.45 or below? \_\_\_\_\_\_\_\_\_\_
* If the actual proportion of juniors at Schueler High School who have a summer job is at least 0.5 (50% of the students), is it unusual to get a random sample of at most 0.45 of juniors who have a summer job?

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

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

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

* Using Tomas’ sample proportion and the results from your simulation, calculate a margin of error for the actual population proportion of juniors at Schueler High School that have a summer job.

Sample proportion ± 2 ∙ std Dev

(where std Dev = std Dev of the simulated sampling distribution.)

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

* Write a sentence interpreting the margin of error in the context of the problem.

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

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

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

* Based on the simulation, which statement below offers the best conclusion?

\_\_\_\_\_ A. Tomas’ sample proportion provides evidence that the principal’s claim was incorrect. It is likely that less than 0.5 (50%) of juniors at Schueler High School have a summer job.

\_\_\_\_\_ B. Tomas’ sample proportion does not provide evidence that the principal’s claim was incorrect. It is possible that at least 0.5 (50%) of juniors at Schueler High School have a summer job.

**Practice Problems**

Problem 1 School Spirit

Lincoln High School is planning a Spirit Day on Friday. Dr. Hernandez, the principal of Lincoln High School, tells members of the student council that he is concerned that less than 40% of the student body will wear school colors on Friday. Jillian, the student council president, is convinced that a much larger percentage of the students will come dressed in school colors. She decides to take a simple random sample of 25 of her classmates and finds that 16 of the 25 students plan on wearing school colors on Friday.

1. What proportion of the students in Jillian’s sample is planning on wearing school colors on Friday?
2. Describe how each of the following might be used in a simulation to generate a sampling distribution.
* Colored chips in a paper bag.
* A random digits table

Jillian uses an online applet to generate a sampling distribution of 80 random samples of size 25 from a population with a proportion of 0.4. The sampling distribution and its statistics are shown below.



1. Based on Jillian’s random sample and the sampling distribution above, what is the margin of error for the true population proportion of Lincoln High School students who plan on wearing school colors on Friday? (Express answers as proportion ± margin of error.)
2. Write a sentence interpreting the margin of error in context.
3. If Jillian decided to take a random sample of size 50, how would this change her margin of error?

Problem 2 Age of Crawling

A child development research study concluded that the mean age, in weeks, that a baby first crawls is on the interval 29.2 weeks to 31.8 weeks.

1. Based on this interval, what is the margin of error?
2. What was the mean crawling age for babies in the sample used in this research study?

Problem 3 National Immunization Survey

In 2010, a National Immunization Survey was conducted. The table on the next page reports the estimated vaccination coverage among children 19 – 35 months of age by state and local area. Results are reported as a percentage ± a margin of error.

1. Row one of the table shows the following information.

|  |  |
| --- | --- |
| **US National** | 74.9±1.2 |

* What is 74.9%, and how was it likely calculated?
* Write a sentence that interprets 74.9±1.2 in the context of this situation.
1. Is there evidence to conclude that the actual percentage of children ages 19 -35 months who have been vaccinated and live in the City of Chicago is higher than the actual percentage of children ages 19 -35 months who have been vaccinated and live in the rest of the State of Illinois? Support your answer with statistical reasoning.
2. Is there evidence that the actual percentage of children ages 19 -35 months who have been vaccinated and live in Los Angeles County, CA is higher than the actual percentage of children ages 19 -35 months who have been vaccinated and live in the rest of the state of California? Support your answer with statistical reasoning.
3. Do you think children in Florida in 2010 are more likely or less likely to have been vaccinated than children living in Idaho? Explain.

|  |  |  |  |
| --- | --- | --- | --- |
| State/Region | Total | State Region | Total |
| US National | 74.9±1.2 | New Hampshire | 84.1±6.1 |
| Alabama | 77.3±5.1 | New Jersey | 66.4±6.8 |
| Alaska | 70.2±6.2 | New Mexico | 70.9±5.9 |
| Arizona | 76.3±5.9 | New York | 69.1±5.3 |
| Arkansas | 79.3±5.5 | NY-City of New York | 65.1±7.8 |
| California | 71.3±6.2 | NY-Rest of State | 73.0±7.1 |
| CA-Los Angeles County | 80.0±5.6 | North Carolina | 77.0±5.2 |
| CA-Rest of State | 68.1±8.2 | North Dakota | 76.0±5.6 |
| Colorado | 71.3±6.1 | Ohio | 76.0±6.6 |
| Connecticut | 75.7±6.8 | Oklahoma | 70.3±6.3 |
| Delaware | 72.9±5.6 | Oregon | 69.3±6.3 |
| Dist. of Columbia | 81.2±6.0 | Pennsylvania | 78.8±4.5 |
| Florida | 85.8±5.1 | PA-Philadelphia County | 74.3±6.1 |
| Georgia | 73.9±6.0 | PA-Rest of State | 79.6±5.2 |
| Hawaii | 76.0±5.8 | Rhode Island | 75.3±6.9 |
| Idaho | 61.2±6.7 | South Carolina | 77.7±5.3 |
| Illinois | 75.9±4.8 | South Dakota | 73.2±6.1 |
| IL-City of Chicago | 76.5±5.3 | Tennessee | 82.3±5.1 |
| IL-Rest of State | 75.7±6.2 | Texas | 74.8±3.9 |
| Indiana | 73.9±5.6 | TX-Bexar County | 78.4±6.0 |
| Iowa | 77.3±6.0 | TX-City of Houston | 74.5±6.1 |
| Kansas | 77.6±6.5 | TX-Dallas County | 72.6±6.6 |
| Kentucky | 72.5±6.2 | TX-El Paso County | 72.6±6.8 |
| Louisiana | 73.8±6.8 | TX-Rest of State | 74.9±5.6 |
| Maine | 70.4±7.3 | Utah | 70.6±6.4 |
| Maryland | 73.3±6.8 | Vermont | 71.0±6.5 |
| Massachusetts | 79.9±5.9 | Virginia | 74.2±6.0 |
| Michigan | 83.4±5.0 | Washington | 73.7±5.4 |
| Minnesota | 74.3±6.5 | WA-Eastern WA | 79.1±6.1 |
| Mississippi | 79.3±4.9 | WA-Western WA | 72.0±6.8 |
| Missouri | 70.3±5.9 | West Virginia | 73.0±5.7 |
| Montana | 64.3±6.8 | Wisconsin | 81.7±5.3 |
| Nebraska | 78.9±5.9 | Wyoming | 67.5±7.0 |
| Nevada | 66.6±6.9 | U.S. Virgin Islands | 48.5±8.2 |

The date in the table above and the following information were taken from:

<http://kff.org/other/state-indicator/percent-who-are-immunized/> 

**Teacher Notes**

New Illinois Learning Standards addressed by this activity:

Content Standards

|  |  |  |
| --- | --- | --- |
|   Major Focus | 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.

**Acknowledgements**

Beth Chance and Allan Rossman have given permission for their applets to be shared with Illinois math teachers. A coin tossing applet, which can be used to simulate taking a single sample of 40 coin tosses or the long run behavior for multiple samples of size 40, is available at <http://www.rossmanchance.com/applets/OneProp/OneProp.htm>.

Consider using this applet to demonstrate tossing a penny 40 times so that students can visually see what is being simulated. Then ask students to use their calculators to generate their own random samples. At the conclusion of the activity, this applet can again be used to show the long run behavior as more and more random samples of size 40 are drawn.



**Answer Key**

Summer Job

* The population is all juniors at Schueler High School.
* The population characteristic of interest is the proportion (percentage) of juniors that have a summer job.
* The sample is the 40 students randomly selected by Tomas.
* The sample statistic is the sample proportion. The sample proportion equals $\frac{18}{40}= 0.45$.
* Describe a simulation carried out using colored chips:

(Answers will vary. Encourage students to write a paragraph.)

A student response should:

* Indicate that a particular color, for example red, will represent juniors who have a summer job. All other colors will represent students who do not have a summer job.
* Indicate that 50% of the chips will be red to correspond to the principal’s claimed value. We are testing a claim about a population – our goal is to create a simulated population that has the claimed value of this population characteristic.
* Describe a method of sampling randomly from the chip population. If the total chip population is small, sampling with replacement should be used.
* Indicate that a random sample of size 30 will be selected from the chip population. The simulated sample must be the same size as Tomas’ sample.
* Describe a simulation using the toss of a coin.

(Answers will vary. Encourage students to write a paragraph.)

A student response should:

* Indicate that a particular outcome, for example heads, will represent juniors who have a summer job. Tails will represent juniors who do not have a summer job.
* Indicate that the probability of a tossed coin landing with the head side up is 0.5. This corresponds to the principal’s claimed value of 0.5 (50%) of the juniors having a summer job.
* Describe a process of tossing the coin and noting the outcome: heads or tails. The collection of outcomes will be the simulated random sample. The proportion of heads in the sample will be the simulated sample proportion.
* Indicate that the coin will be tossed 30 times to generate a sample of size 30. The simulated sample must be the same size as Tomas’ sample.

Analysis

* Describe the shape center and spread of the distribution of the Simulated *Sample Proportions.*

(Answers will vary.)

The distribution (dot plot) of the proportions from our simulated samples is approximately symmetrical. The center of the distribution is about 0.5. (The mean of all of the sample proportions is approximately 0.5.) The sample proportions vary from a low of 0.325 to a high of 0.675. A typical deviation in the sample proportions from the mean proportion of 0.5 is 0.075 (the standard deviation).

* 0.5
* (Answers will vary.)

Suggested response: No, I am not surprised that the center of the distribution is 0.5. We set the simulated population proportion to be 0.5. The center of the distribution of all the sample proportions should be equal to this population proportion.

* The mean should be close to 0.5 and the standard deviation should be close to 0.079.

.

.

.

* If the actual proportion of juniors at Schueler High School who have a summer job is at least 0.5 (50% of the juniors), is it unusual to get a random sample of at most 0.45 juniors who have a summer job?

No, it is not unusual to get a random sample of 0.45 or less if the population proportion is 0.5. Many of our random samples had a proportion of 0.45 or less. It is a likely outcome when the population proportion is 0.5.

* (Answers will vary based on the value of the class’ sampling distribution standard deviation.)

Approximate answer: 0.45 ± 2 ∙ 0.079

 0.45 ± 0.158

* Based on Tomas’ random sample and the standard deviation from our simulated sampling distribution, the actual population proportion of juniors at Schueler High School who have a summer job is on the interval of 0.292 to 0.608. (Notice that Mrs. Johnson’s claimed value of at least 0.5 (50%) falls on this interval. )
* B

Practice Problems

Problem 1 School Spirit

1. $\frac{16}{25}$ or 0.64 is the proportion of students in Jillian’s sample who plan on wearing school colors on Friday.
* Describe a simulation carried out using colored chips:

(Answers will vary. Encourage students to write a paragraph.)

A student response should:

* Indicate that a particular color, for example red, will represent students who plan on wearing school colors on Friday. All other colors will represent all other students.
* Indicate that 40% of the chips will be red to correspond to the principal’s claimed value.
* Describe a method of sampling randomly from the chip population.
* Indicate that a random sample of size 25 will be selected from the chip population. The simulated sample must be the same size as Jillian’s sample.
* Describe a simulation using a random digits table.

(Answers will vary. Encourage students to write a paragraph.)

A student response should cover all the points mentioned above. Here is a sample response:

I will let the numbers 1, 2, 3, and 4 represent students who plan on wearing school colors on Friday. The numbers 5 – 9 and 0 will represent all other students. The probability of selecting a 1 – 4 is 0.4 and corresponds to the Dr. Hernandez’s claimed population proportion of 0.4. My sample must be the same size as Jillian’s sample. I will select a row in the random digits table and then pick the first 25 numbers. These 25 numbers will represent my sample of 25 students. Next I will determine the proportion of 1, 2, 3 and 4 in my sample of 25. This will be my simulated sample proportion.

1. Margin of error using Jillian’s sample and the standard deviation from the given sampling distribution: proportion ± 2 ∙ std Dev

 0.64 ± 2 ∙ 0.098

 0.64 ± 0.196

1. The actual population proportion of students at Lincoln High School who plan on wearing school colors on Friday is on the interval of 0.444 to 0.836. Notice that Dr. Hernandez’s claimed value of 40% (or less) is not on this interval. There is evidence to conclude that Dr. Hernandez’s claim is incorrect. The actual population proportion of students who plan on wearing school colors is higher.

If Jillian increased her sample size to 50, the margin of error would be narrower. A larger sample size produces a narrower sampling distribution – less spread – and a smaller standard deviation.

Problem 2 Age of Crawling

1. The margin of error is half of the length of the interval.

31.8 – 29.2 = 2.6.

$$\frac{1}{2} \left( 2.6\right)=1.3$$

The margin of error is ± 1.3 weeks

1. The mean age of crawling is the midpoint of the interval.

$\frac{29.2+31.8}{2}=30.5$ weeks

The mean age at which babies in the sample begin to crawl is 30.5 weeks.

Problem 3 National Immunization Survey

1. 74.9% or .749 is the proportion of children ages 19 – 35 months who have been vaccinated in a national sample of children ages 19 – 35 months. The table and accompanying information do not indicate if the sample was a random sample. With more research, it is possible that we would find that the samples used to generate the table of information were random samples of the population.
2. The margin of error for children ages 19 – 35 months that have been vaccinated and live in the City of Chicago is 76.5 ± 5.3. Assuming that the sample that was used was randomly selected, then the actual population percentage of children in this age range who live in Chicago that were vaccinated is on the interval 71.2 % to 81.8%. The margin of error for the rest of Illinois is 75.7 ± 6.2. The actual population percentage of children ages 19 -35 months who live in other areas of the State of Illinois is on the interval 69.5% to 81.95%. These two intervals overlap. There is not evidence that the actual population percentage of children in this age group who live in Chicago have a higher vaccination rate than those who live in the rest of the State of Illinois. The overlapping intervals indicate that it is possible that Chicago’s percentage could be higher, lower or even equal to the population percentage for the rest of Illinois.
3. The margin of error for children ages 19-35 months that have been vaccinated and live in Los Angeles County is 80.0 ± 5.6. Assuming that the sample was randomly selected (and therefore representative), the actual population percentage of children ages 19-35 that have been vaccinated and live in Los Angeles County is 74.4% to 85.6%. The margin of error for the rest of the State of California is 68.1 ± 8.2. The actual population percentage of children in this age range from the rest of California who have been vaccinated is 59.9% to 76.3%. While there is less overlap in these intervals than in (b), the intervals do still overlap. Since the actual population percentages can fall anywhere on the stated intervals, it is possible that the population percentage of children ages 19 – 35 months in Los Angeles County could be more, less or equal to the population percentage for the rest of the State of California.
4. The margin of error for children ages 19 - 35 months living in Florida in 2010 who were vaccinated is 85.8 ± 5.1. The actual population percentage of Florida children in the given age range who have been vaccinated is on the interval 80.7% to 90.0%. The actual population percentage for Idaho children in the same age range has a margin of error of 61.2 ± 6.7 and falls on an interval of 54.5% to 67.9%. These two intervals do not overlap. The actual percentage of vaccinated of Idaho children in 2010 is lower than the actual percentage of Florida children. Florida children ages 19 – 35 months were more likely to have been vaccinated.