## Sampling Presidents

One of the most important ideas in statistics is that a lot can be learned about a population by studying a sample. Consider the population of all 44 presidents of the United States.

Suppose that you want to estimate the average (mean) height of presidents of the United States.

You consider selecting a sample of five presidents to estimate this value. Suppose that you decide to select a sample of five presidents by asking a classmate to name five presidents that he or she is familiar with.

1. Is there any reason to suspect that this sampling method would be biased, with respect to estimating the average height of the 44 presidents of the United States? Explain your answer.

**Collecting Samples**

To examine whether this sampling method is biased, you will collect some data.

Think of five names of presidents that you would consider representative of all the presidents of the United States.

1. Walk around the room and collect five names of presidents from five classmates. Record the names in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Student Number** | **President #1** | **President #2** | **President #3** | **President #4** | **President #5** |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |

1. For each of the five samples, determine the height for each president named, and find the average height for each sample. (Note: The *Presidents* data is located on pages 20 and 21 at the end of this activity)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Student Number** | **President #1 Height** | **President #2 Height** | **President #3 Height** | **President #4 Height** | **President #5 Height** | **Average Height** |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |

Record your five sample means on the dotplot on the board.

1. Sketch the plot of all of the sample estimates. Make sure to label the axis appropriately.
2. Are these five sample means parameters or statistics? Explain your answer.
3. The actual population average height based on all 44 presidents is 70.73 inches. Where does this value fall in the above plot? Were most of the sample estimates *around* the population mean? Explain.
4. For how many samples did the sample estimate exceed the population average? What proportion of the samples is this?
5. Based on your answer to question 7, is a sample estimate just as likely to be above the population average as it is to be below the population average?
6. Based on the plot, is asking five classmates to name five presidents a reasonable (unbiased) way to estimate the average height for all 44 presidents? Explain.

When the sampling method produces characteristics of the sample that systematically differ from those characteristics of the population, you say that the sampling method is *biased*. To try to eliminate potential biases, it is better to take a *random* sample. This should create a representative sample, no matter what variable is focused on. Humans are not very good “random samplers,” so it is important to use other techniques to do the sampling for us.

**SIMPLE RANDOM SAMPLING**

The first step in drawing a simple random sample is to obtain a *sampling frame* or a list of each member of the population. Then, you can use software to randomly select a sample from the sampling frame.

* Go to <http://www.rossmanchance.com/applets/>.
* Click on *Sampling from a Finite Population*, under the *Sampling Distribution Simulations* section.
* You will need to enter the data into the applet. To do this, open the *Presidents.csv* file. Copy the 44 values from the Height column (NOT including the header row).
* In the applet, click on Paste Numeric Data. Paste the 44 values into the window that pops up. Then click Ok.

If the data has been entered correctly, the population distribution for the height will be displayed in the graph window on the right. To double-check your data entry, make sure that μ = 70.73 and σ = 2.70 in the upper-right-hand corner of the plot.

* To draw a simple random sample of size 5 from the population of presidents, type the value 5 in the box next to sample size(on the left side of the page) and then click Draw Samples. The random sample of size 5 you drew will appear in the plot at the bottom of the webpage (note: need to scroll down to see the plot).

1. What is the value from your simple random sample?

Record your simple random sample mean on the dotplot on the board.

1. Sketch the plot of all of the sample estimates. Make sure to label the axis appropriately.
2. This time how many samples had a sample average that was larger than the population average? What proportion of the statistics is this?
3. If the sampling method is unbiased the estimates of the population average should be centered ‘around’ which value?
4. Does this appear that random sampling is an unbiased method?

To really examine the long-term patterns of this sampling method, you will use technology to take many, many samples.

* Click the Reset button.
* Type in 5 in the box next to sample size and change the Number of samplesvalue from 1 to 100.
* Click Draw Samples.

1. Describe what the applet is doing when the number of samples values is changed from 1 to 100.

For each sample, the sample mean is calculated, and this mean is plotted in the plot at the bottom of the webpage. The red arrow indicates the average of the 100 sample averages (or sample means).

1. Record the value of the average of the 100 sample averages.
2. Does the sampling method appear to be an unbiased sampling method? Explain.

* Change the sample size from 5 to 20. Make sure that Number of samplesis setto 100.
* Click Draw Samples.

1. Record the value of the average of the 100 sample averages, when sample size is 20.
2. Does the sampling method still appear to be unbiased? Explain.
3. What has changed about the distribution of the sample averages? Why does this make sense?

| **Number** | **President** | **Height**  **(in inches)** | **Age at Inauguration**  **(in years)** | **Political Party** |
| --- | --- | --- | --- | --- |
| 1 | George Washington | 73.5 | 57 | No party |
| 2 | John Adams | 67 | 61 | Federalist |
| 3 | Thomas Jefferson | 74.5 | 57 | Democratic-  Republican |
| 4 | James Madison | 64 | 57 | Democratic-  Republican |
| 5 | James Monroe | 72 | 58 | Democratic-  Republican |
| 6 | John Quincy Adams | 67.5 | 57 | Democratic-  Republican |
| 7 | Andrew Jackson | 73 | 61 | Democratic |
| 8 | Martin Van Buren | 66 | 54 | Democratic |
| 9 | William Henry Harrison | 68 | 68 | Whig |
| 10 | John Tyler | 72 | 51 | No party |
| 11 | James K. Polk | 68 | 49 | Democratic |
| 12 | Zachary Taylor | 68 | 64 | Whig |
| 13 | Millard Fillmore | 69 | 50 | Whig |
| 14 | Franklin Pierce | 70 | 48 | Democratic |
| 15 | James Buchanan | 72 | 65 | Democratic |
| 16 | Abraham Lincoln | 76 | 52 | Republican |
| 17 | Andrew Johnson | 70 | 56 | Democratic |
| 18 | Ulysses S. Grant | 68 | 46 | Republican |
| 19 | Rutherford B. Hayes | 68.5 | 54 | Republican |
| 20 | James A. Garfield | 72 | 49 | Republican |
| 21 | Chester A. Arthur | 72 | 51 | Republican |
| 22 | Grover Cleveland | 71 | 47 | Democratic |
| 23 | Benjamin Harrison | 66 | 55 | Republican |
| 24 | Grover Cleveland | 71 | 55 | Democratic |
| 25 | William McKinley | 67 | 54 | Republican |
| 26 | Theodore Roosevelt | 70 | 42 | Republican |
| 27 | William Howard Taft | 71.5 | 51 | Republican |
| 28 | Woodrow Wilson | 71 | 56 | Democratic |
| 29 | Warren G. Harding | 72 | 55 | Republican |
| 30 | Calvin Coolidge | 70 | 51 | Republican |
| 31 | Herbert Hoover | 71.5 | 54 | Republican |
| 32 | Franklin D. Roosevelt | 74 | 51 | Democratic |
| 33 | Harry S Truman | 69 | 60 | Democratic |
| 34 | Dwight D. Eisenhower | 70.5 | 62 | Republican |
| 35 | John F. Kennedy | 72 | 43 | Democratic |
| 36 | Lyndon B. Johnson | 76 | 55 | Democratic |
| 37 | Richard Nixon | 71.5 | 56 | Republican |
| 38 | Gerald Ford | 72 | 61 | Republican |
| 39 | Jimmy Carter | 69.5 | 52 | Democratic |
| 40 | Ronald Reagan | 73 | 69 | Republican |
| 41 | George H. W. Bush | 74 | 64 | Republican |
| 42 | Bill Clinton | 74 | 46 | Democratic |
| 43 | George W. Bush | 71.5 | 54 | Republican |
| 44 | Barack Obama | 73 | 47 | Democratic |