

Lincoln and Mandela: Random Sampling and Random Assignment

Allan J. Rossman

Department of Statistics, Cal Poly – San Luis Obispo

Email: arossman@calpoly.edu Twitter: [@allanjrossman](https://twitter.com/allanjrossman)

Blog: <https://askgoodquestions.blog>

presented for “Chat & Chew” at AMATYC Conference in Milwaukee on November 14, 2019

Activity 1: Lincoln

a) Select a sample by circling ten words from this population of 268 words:

Four score and seven years ago, our fathers brought forth upon this continent a new nation: conceived in liberty, and dedicated to the proposition that all men are created equal.

Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battlefield of that war.

We have come to dedicate a portion of that field as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this.

But, in a larger sense, we cannot dedicate, we cannot consecrate, we cannot hallow this ground. The brave men, living and dead, who struggled here have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember, what we say here, but it can never forget what they did here.

It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us, that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion, that we here highly resolve that these dead shall not have died in vain, that this nation, under God, shall have a new birth of freedom, and that government of the people, by the people, for the people, shall not perish from the earth.

b) For each word in your sample, record how many letters are in the word.

c) Calculate the average (mean) number of letters per word in your sample. [*Hint*: Add up the number of letters in each word and divide by ten.]

d) Combine your sample average with those of your classmates to produce a dotplot of sample averages below. Be sure to label the horizontal axis appropriately.



e) The average number of letters per word in the population of all 268 words is 4.295. Mark this value on the dotplot.

f) Would you say that this sampling method (asking people to simply circle ten representative words) is biased? If so, in which direction? Explain how you can tell from the dotplot.

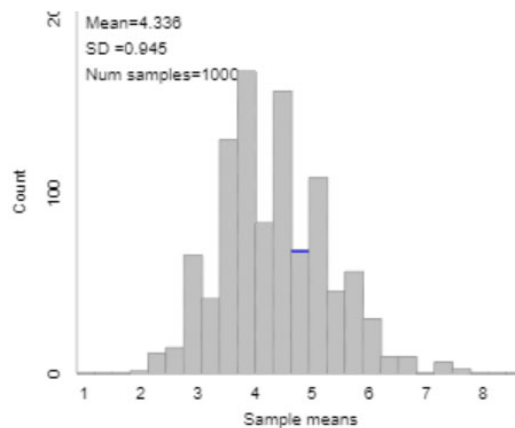
g) Suggest some reasons why this sampling method turned out to be biased as it did.

h) Would using this same sampling method but with a larger sample size (say, asking you to circle 20 words) reduce or eliminate the sampling bias? Explain.

i) Consider a different sampling method: close your eyes and point to the page ten times to select the words for your sample. Would this sampling method be biased or unbiased? Explain.

j) Suggest how you might employ a different sampling method that would be unbiased.

An applet (www.rossmanchance.com/applets/OneSample.html?population=gettysburg) selected 1000 random samples of size 5 words each, producing the following graph of sample means:



k) Is random sampling biased or unbiased? What aspect of the graph addresses this?

l) How would the distribution of sample means change if the sample size had been 20 rather than 5? Address center, variability, and shape.

m) Is taking a large sample *always* preferable to taking a smaller sample? Explain. (*Hint: Remember the Literary Digest poll.*)

Ask Good Questions: a weekly blog that provides ideas, examples, activities, assessments, and advice for teaching introductory statistics based on a three-word teaching philosophy. (<https://askgoodquestions.blog>)

Lincoln and Mandela: Random Sampling and Random Assignment

Allan J. Rossman

Department of Statistics, Cal Poly – San Luis Obispo

Email: arossman@calpoly.edu Twitter: [@allanjrossman](https://twitter.com/allanjrossman)

Blog: <https://askgoodquestions.blog>

presented for “Chat & Chew” at AMATYC Conference in Milwaukee on November 14, 2019

Activity 2: Mandela (Version A)

a) Answer these two questions:

- Was Nelson Mandela, first president of South Africa following apartheid, younger or older than 16 years old when he died?
- Make a guess for Nelson’s Madela’s age at his death.

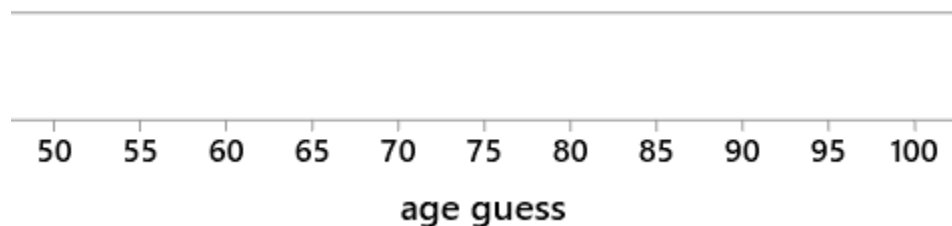
b) Identify the observational units and variables in this study.

c) Was this an observational study or a randomized experiment? Explain.

d) Did this study make use of random sampling, random assignment, both, or neither? Explain the purpose of any randomness that was used in this study.

e) Some people know more about world history than others. Is knowledge of world history a confounding variable in this study? If not, how did we control for this?

f) Produce dotplots of the data on the axes below:



g) Describe what the data reveal about the research question that motivated the data collection.

h) If the age guesses between the two groups differ significantly, would it be appropriate to draw a cause-and-effect conclusion? Between what and what? Explain.

i) To what population is it reasonable to generalize the results of this study? Explain.

- **Random sampling** concerns how to select observational units for a sample. Random sampling allows for *generalizing* the results of a sample to the larger population.
- **Random assignment** pertains to how observational units come to be in groups to be compared. Random assignment allows for the possibility of drawing a *cause-and-effect* conclusion.

Lincoln and Mandela: Random Sampling and Random Assignment

Allan J. Rossman

Department of Statistics, Cal Poly – San Luis Obispo

Email: arossman@calpoly.edu Twitter: [@allanjrossman](https://twitter.com/allanjrossman)

Blog: <https://askgoodquestions.blog>

presented for “Chat & Chew” at AMATYC Conference in Milwaukee on November 14, 2019

Activity 2: Mandela (Version B)

a) Answer these two questions:

- Was Nelson Mandela, first president of South Africa following apartheid, younger or older than 160 years old when he died?
- Make a guess for Nelson’s Madela’s age at his death.

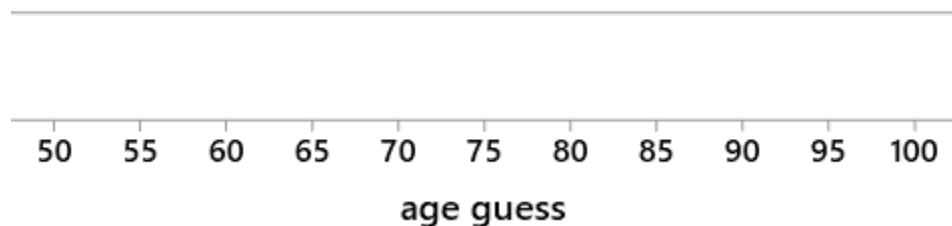
b) Identify the observational units and variables in this study.

c) Was this an observational study or a randomized experiment? Explain.

d) Did this study make use of random sampling, random assignment, both, or neither? Explain the purpose of any randomness that was used in this study.

e) Some people know more about world history than others. Is knowledge of world history a confounding variable in this study? If not, how did we control for this?

f) Produce dotplots of the data on the axes below:



g) Describe what the data reveal about the research question that motivated the data collection.

h) If the age guesses between the two groups differ significantly, would it be appropriate to draw a cause-and-effect conclusion? Between what and what? Explain.

i) To what population is it reasonable to generalize the results of this study? Explain.

- **Random sampling** concerns how to select observational units for a sample. Random sampling allows for *generalizing* the results of a sample to the larger population.
- **Random assignment** pertains to how observational units come to be in groups to be compared. Random assignment allows for the possibility of drawing a *cause-and-effect* conclusion.