

Optional Sampling Distribution Class Activity 1

Exploring Sampling Variability for Mean Averages with a Sampling Distribution

The goal of this activity is to explore how well random samples approximate population values. Normally we do not know population values and we must use a sample value to approximate the population value. This is called a “point estimate”. For this activity, we will look at some population data from International Coffee Organization (ICO). We will be using the “Columbian Mild” price data in U.S. cents per pound. The population mean average price was 136.43 cents per pound. Again, in real data analysis, we often do not know the population value, but for this activity, it is useful for comparison purposes.

Open the “Sampling Distribution Data 1” in Excel. A total of 120 random samples have been taken from the Columbian Mild data. All the data sets have 30 coffee prices. Each person in the class will be finding the mean of a few of these data sets. Once you find your sample means, you will put a magnet up or draw a dot on the board to represent the sample mean you found. When everyone’s magnets or dots are up on the board, we will have generated a “sampling distribution”.

Answer the following questions:

1. The population mean was 136.43 cents. How many cents was the sample mean you calculated from the population mean of 136.43 cents? (If you calculated more than one sample mean, answer the question for all the sample means you calculated.) This is called the “Margin of Error”.
 2. Look at the dots or magnets on the board. Did all the sample means come out to be the same as the population mean of 136.43cents? Why do you think this happened? Aren’t random samples supposed to be good approximations of the population? What does this tell you about sampling variability?
 3. Normally, we may have only one random sample. If all you knew was one of the random samples on the board, how difficult would it be to determine that the population mean is really 136.43 cents? What does this tell us about the difficulty in determining population values from one random sample?
 4. Estimate the shape and center of the sampling distribution on the board. Is the center of the graph close to the population mean of 136.43? Would the center of the sampling distribution be a better approximation of the population mean than a single sample mean?
 5. The standard deviation of a sampling distribution is often called the “standard error” and is an important part of inferential statistics. Estimate how far typical dots are from the center of the sampling distribution. This is the standard deviation of the sampling distribution, which is called “Standard Error”.
-

Optional Sampling Distribution Class Activity 2

Exploring Sampling Variability for Percentages with a Sampling Distribution

The goal of this activity is to explore how well random sample percentages approximate population percentages. Normally we do not know population percentage and we must use a sample percentage to approximate the population percentage. This is called a “point estimate”. For this activity, we will be flipping coins 30 times and count the number of tails. Then calculate the sample percentage of tails. Each person will do three sets of 30 and therefore get three sample percentages. Again, in real data analysis, we often do not know the population value, but for this activity, it is useful for comparison purposes. Our goal is to see how well random sample percentages approximate population percentages.

Each person in the class will be finding three sample percentages. Once you find each sample percent, you will put a magnet up or draw a dot on the board to represent the sample percent you found. When everyone’s magnets or dots are up on the board, we will have generated a “sampling distribution” of sample percentages.



Answer the following questions:

1. In a perfect world and a fair coin, what should the population percentage for getting tails be? So in a sample of 30 how many times do we expect to get tails? In sampling, we often do not get what we expect. How far were the sample percentages you calculated from the population percentage?
 2. Look at the dots or magnets on the board. Did all the sample percentages come out to be the same as the population percentage? Why do you think this happened? Aren't random samples supposed to be good approximations of the population? What does this tell you about sampling variability?
 3. Normally, we may have only one random sample. If all you knew was one of the sample percentage on the board, and you never knew the expected population value, how difficult would it be to determine what the population percentage really is? What does this tell us about the difficulty in determining population values from one random sample?
 4. Estimate the shape and center of the sampling distribution on the board. Is the center of the graph close to the population percentage of 0.5? Would the center of the sampling distribution be a better approximation of the population percentage than a single sample percentage?
 5. The standard deviation of a sampling distribution is often called the "standard error" and is an important part of inferential statistics. Estimate how far typical dots are from the center of the sampling distribution. This is the standard deviation of the sampling distribution, which is called "Standard Error".
-

Problem Set Section 2B

Directions: Answer the following questions about sampling distributions.

1. Describe the process of making a sampling distribution.
2. What can sampling distributions tell us about sampling variability?
3. What is a point estimate? Discuss how point estimates create confusion for people reading articles and scientific reports.
4. Discuss the shape of sampling distributions. When the population is skewed, is the sampling distribution always normal? Why is it important for a sampling distribution to be normal? In the examples in this section, which statistics had a normal sampling distribution? Which statistics had a skewed sampling distribution?
5. Explain how the standard error is calculated. What does the standard error tell us about sample statistics and the population parameter? Why is the standard error only accurate when the sampling distribution is normal?
6. What is the difference between standard error and margin of error? Is the standard error smaller or larger than the margin of error?

(#7-16) For the following problems, copy the indicated census data set from the Math 140 Survey Data at www.matt-teachout.org. We will be assuming this is an unbiased census and therefore know the population mean. Open StatKey at www.lock5stat.com. Under the "sampling distributions" menu, click on "mean". You should see "sampling distribution for the mean". Under "edit data" paste in the indicated data set. Under "choose samples of size n", put in the indicated sample size. Create a sampling distribution and then answer the following questions.

7. Use StatKey to create a sampling distribution with sample size 10 from the Age in Years census data (Math 140 Survey Data).
 - a) What was the shape and mean average of the population?
 - b) Were all the sample means the same as the population mean?
 - c) Were all the sample means the same as each other?
 - d) How many random samples did you take when you created the sampling distribution?
 - e) What is the shape of the sampling distribution?



- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.

8. Use StatKey to create a sampling distribution with sample size 100 from the Age in Years census data (Math 140 Survey Data)

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.
- h) How does the standard error for sample size 10 compare to the standard error for sample size 100?
- i) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?

9. Use StatKey to create a sampling distribution with sample size 10 from the sleep hours per night census data (Math 140 Survey Data)

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.

10. Use StatKey to create a sampling distribution with sample size 25 from the sleep hours per night census data (Math 140 Survey Data)

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.
- h) How does the standard error for sample size 10 compare to the standard error for sample size 25?
- i) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 25?

11. Use StatKey to create a sampling distribution with sample size 10 from the cell phone bill (in dollars per month) census data (Math 140 Survey Data).

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.



12. Use StatKey to create a sampling distribution with sample size 100 from the cell phone bill (in dollars per month) census data (Math 140 Survey Data).

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.
- h) How does the standard error for sample size 10 compare to the standard error for sample size 100?
- i) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?

13. Use StatKey to create a sampling distribution with sample size 10 from the travel time to get to school in minutes (Math 140 Survey Data).

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.

14. Use StatKey to create a sampling distribution with sample size 40 from the travel time to get to school in minutes (Math 140 Survey Data).

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.
- h) How does the standard error for sample size 10 compare to the standard error for sample size 40?
- i) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 40?

15. Use StatKey to create a sampling distribution with sample size 10 from the work hours per week for COC college students (Math 140 Survey Data).

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?
- g) What is the standard error? Write a sentence explaining the meaning of the standard error.

16. Use StatKey to create a sampling distribution with sample size 40 from the work hours per week for COC college students (Math 140 Survey Data).

- a) What was the shape and mean average of the population?
- b) Were all the sample means the same as the population mean?
- c) Were all the sample means the same as each other?
- d) How many random samples did you take when you created the sampling distribution?
- e) What is the shape of the sampling distribution?
- f) What is the center (mean) of the sampling distribution? Is it relatively close to the population mean?



- g) What is the standard error? Write a sentence explaining the meaning of the standard error.
- h) How does the standard error for sample size 10 compare to the standard error for sample size 40?
- i) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 40?

(#17-26) The following population proportions come from the Math 140 Survey Data at www.matt-teachout.org. We will be assuming this is an unbiased census and therefore know the population proportion (%). Open StatKey at www.lock5stat.com. Under the “sampling distributions” menu, click on “proportion”. You should see “sampling distribution for a proportion”. Under “edit proportion”, enter the given population proportion. Create a sampling distribution and then answer the following questions.

17. A census of COC statistics students in the fall 2015 semester indicated that the population proportion of statistics students with brown hair is 0.537. Use this population proportion to create a sampling distribution with sample size 10 with StatKey.

- a) Were all the sample proportions the same as the population proportion?
- b) Were all the sample proportions the same as each other?
- c) How many random samples did you take when you created the sampling distribution?
- d) What is the shape of the sampling distribution?
- e) What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- f) What is the standard error? Write a sentence explaining the meaning of the standard error.

18. A census of COC statistics students in the fall 2015 semester indicated that the population proportion of statistics students with brown hair is 0.537. Use this population proportion to create a sampling distribution with sample size 100 with StatKey.

- a) Were all the sample proportions the same as the population proportion?
- b) Were all the sample proportions the same as each other?
- c) How many random samples did you take when you created the sampling distribution?
- d) What is the shape of the sampling distribution?
- e) What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- f) What is the standard error? Write a sentence explaining the meaning of the standard error.
- g) How does the standard error for sample size 10 compare to the standard error for sample size 100?
- h) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?

19. A census of COC statistics students in the fall 2015 semester indicated that the population proportion of statistics students that smoke cigarettes is 0.091. Use this population proportion to create a sampling distribution with sample size 10 with StatKey.

- a) Were all the sample proportions the same as the population proportion?
- b) Were all the sample proportions the same as each other?
- c) How many random samples did you take when you created the sampling distribution?
- d) What is the shape of the sampling distribution?
- e) What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- f) What is the standard error? Write a sentence explaining the meaning of the standard error.



20. A census of COC statistics students in the fall 2015 semester indicated that the population proportion of statistics students that smoke cigarettes is 0.091. Use this population proportion to create a sampling distribution with sample size 100 with StatKey.

- Were all the sample proportions the same as the population proportion?
- Were all the sample proportions the same as each other?
- How many random samples did you take when you created the sampling distribution?
- What is the shape of the sampling distribution?
- What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- What is the standard error? Write a sentence explaining the meaning of the standard error.
- How does the standard error for sample size 10 compare to the standard error for sample size 100?
- How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?

21. Approximately 60% of college students in the U.S. were able to finish their bachelor's degree in six years. Use this population proportion to create a sampling distribution with sample size 10 with StatKey.

- Were all the sample proportions the same as the population proportion?
- Were all the sample proportions the same as each other?
- How many random samples did you take when you created the sampling distribution?
- What is the shape of the sampling distribution?
- What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- What is the standard error? Write a sentence explaining the meaning of the standard error.

22. Approximately 60% of college students in the U.S. were able to finish their bachelor's degree in six years. Use this population proportion to create a sampling distribution with sample size 100 with StatKey.

- Were all the sample proportions the same as the population proportion?
- Were all the sample proportions the same as each other?
- How many random samples did you take when you created the sampling distribution?
- What is the shape of the sampling distribution?
- What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- What is the standard error? Write a sentence explaining the meaning of the standard error.
- How does the standard error for sample size 10 compare to the standard error for sample size 100?
- How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?

23. Approximately 9.4% of all adults in the U.S. have diabetes. Use this population proportion to create a sampling distribution with sample size 10 with StatKey.

- Were all the sample proportions the same as the population proportion?
- Were all the sample proportions the same as each other?
- How many random samples did you take when you created the sampling distribution?
- What is the shape of the sampling distribution?
- What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- What is the standard error? Write a sentence explaining the meaning of the standard error.

24. Approximately 9.4% of all adults in the U.S. have diabetes. Use this population proportion to create a sampling distribution with sample size 100 with StatKey.

- Were all the sample proportions the same as the population proportion?
- Were all the sample proportions the same as each other?
- How many random samples did you take when you created the sampling distribution?
- What is the shape of the sampling distribution?



- e) What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- f) What is the standard error? Write a sentence explaining the meaning of the standard error.
- g) How does the standard error for sample size 10 compare to the standard error for sample size 100?
- h) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?

25. Approximately 90% of all lung cancer cases are caused by cigarette smoking. Use this population proportion to create a sampling distribution with sample size 10 with StatKey.

- a) Were all the sample proportions the same as the population proportion?
- b) Were all the sample proportions the same as each other?
- c) How many random samples did you take when you created the sampling distribution?
- d) What is the shape of the sampling distribution?
- e) What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
- f) What is the standard error? Write a sentence explaining the meaning of the standard error.

26. Approximately 90% of all lung cancer cases are caused by cigarette smoking. Use this population proportion to create a sampling distribution with sample size 100 with StatKey.

- a) Were all the sample proportions the same as the population proportion?
 - b) Were all the sample proportions the same as each other?
 - c) How many random samples did you take when you created the sampling distribution?
 - d) What is the shape of the sampling distribution?
 - e) What is the center (mean) of all the sample proportions in the sampling distribution? Is it relatively close to the population proportion (π)?
 - f) What is the standard error? Write a sentence explaining the meaning of the standard error.
 - g) How does the standard error for sample size 10 compare to the standard error for sample size 100?
 - h) How does the shape of the sampling distribution for sample size 10 compare to the shape of the sampling distribution for sample size 100?
-

