Sampling Distributions Lesson Plan

Tiffany Timbers

After this lesson students will be able to:

- Define a population and sampling distribution
- Describe the central limit theorem with regards to the relationship between the population and sampling distributions
- Explain how sample size effects a sampling distribution's standard deviation and shape
- Given a population mean and standard deviation, as well as a sample size, predict the probability of observing values less than, or greater than a given cutoff value.

Required materials

- Whiteboard
- Overhead projector
- 3 decks of playing cards
- Laptop (for tutor and each student) with internet access

Lesson (55 minutes)

1. Motivation (2.5 minutes)

- Sampling distributions are the basis for inferential statistics
- For example, a solid understanding of sampling distributions will be extremely useful when we want to determine if two (or more) samples are derived from the same or different populations

2. Mean of the sampling distribution is equal to the mean of the population (15 minutes)

An interactive and live demonstration will be used to introduce the concepts of population distribution, sample distribution, sampling distribution, population mean, sample mean and central limit theorem, as well as to illustrate the concept that the mean of the sampling distribution is equal to the mean of the population. Three decks of cards will be assembled ahead of time (using only cards of value ace - ten) so that they have this known bimodal distribution which will represent the 'population'. The deck of cards will be passed to each student in class, they will shuffle the deck and deal themselves 5 cards. They will write these numbers down, replace the cards into the deck, and pass the deck onto the next student. Each student will calculate the mean from their cards, and share that value with the class. Each student's sample mean will be added to a histogram on the whiteboard and together a sampling distribution will be generated for the sample mean. This sampling distribution will then be compared to the population distribution and its mean. This will demonstrate graphically that the mean of the sampling distribution is equal to (or at least very close to) the mean of the population.

3. Two other important principles regarding the sampling and population distributions (10 min)

Using an online simulation and a shared Google spreadsheet we will explore the effect of sample size on the sampling distribution. The goal will be to demonstrate that:

- When the sample size is sufficiently large, the shape of the sampling distribution is approximately normal irrespective of the shape of the population distribution.
- The standard deviation of the sampling distribution is equal to the standard deviation of the population distribution divided by the square root of the sample size (n).

Each student will be assigned a sample size number. The student will complete a sampling distribution simulation using the assigned numbers as the sample size and select the following additional parameters: Skewed, Mean and 10000 replications. After completing the simulation, the students will copy and paste the values they obtained for the sampling distribution's standard deviation, skew and kurtosis into the correct cell in the shared Google spreadsheet. At this point the terms skew and kurtosis will be reintroduced (they will have learned these in a previous class). The Google spreadsheet has previously been set-up such that when students paste the values from their simulations into the appropriate cells, the spreadsheet adds those values to scatter plots for sample size versus standard deviation, sample size versus skew, and sample size versus kurtosis. Once the spreadsheet has been populated the class will discuss their observations. The tutor will guide and focus the discussion around how increasing the sample size influences the shape of the sampling distribution and its standard deviation.

4. Socrative quiz (5 min)

A free online quizzing tool, [Socrative], will be used to administer the following multiple choice questions to uncover any misconceptions that may have arisen during the last two exercises. If any are apparent, time will be taken to correct and clarify. Correct answers are highlighted in red.

i. Which of the following are not true:
a) as you increase the sample size, the standard deviation of the sampling distribution will decrease
b) as you increase the sample size, the mean of the sampling distribution approaches the mean of the population distribution
c) the mean of the sampling distribution is equal to the mean of the population distribution
d) as you increase the sample size, the sampling distribution becomes more normally distributed

ii. Which answer(s) below are true about the sample distribution:
a) it closely resembles the population distribution
b) it becomes more normally distributed as you increase the sample size
c) as you increase the sample size, the standard deviation of the sample distribution will decrease

5. Applications of the sampling distribution (5 min)

How to use what we have just learned about the sampling distribution to solve an applied problem (outlined below) will be demonstrated to the class.

You are planning a day of sailing on the ocean on a very large boat for 40 people and you plan to bring 112 L of water. The average human is recommended to drink 2.6 L of water each day (with a standard deviation of 0.8 L). What is the probability that you will run out of water?

6. Group problem solving (15 min)

Students will work in groups of 3 to try to solve the problems below. The answers to the first will be taken up in class. Answers to all will be made available online after class. *If lesson sections 1-5 take longer than anticipated, this activity will be assigned as homework instead of being done in class.*

i. The speeds of vehicles in a school zone have a mean of 36 km/h with a standard deviation of 6 km/h, despite the speed limit being 30 km/h. What is the probability that the mean speed of the next 20 cars will be less than the speed limit?

ii. The mean gas bill for all detached homes in Narniaville is $65 a month with a standard deviation of $15. What is the probability that a random sample of gas bills from 50 households in Narniaville will be more than the population mean by at least $5.

iii. A Steel Mill produces rebar with a mean length of 3m and a standard deviation of 0.01m. During the quality control process an inspector measures the length of 10 randomly sampled pieces of rebar to determine if the machine that cuts the rebar needs to be recalibrated. The inspector makes this call if the mean of the sampled rebar is less than 2.995m or greater than 3.05m. What is the probability that it will be decided that the cutting machine needs to be recalibrated after the next inspection?

7. Wrap-up (2.5 min)

A student driven concept map summarizing what was learned will be drawn on the whiteboard. A photo will be taken of the concept map and posted to the course website.