Lesson 17: Sampling Variability

Classwork

Example 1: Estimating a Population Mean

The owners of a gym have been keeping track of how long each person spends at the gym. Eight hundred of these times (in minutes) are shown in the population tables located at the end of the lesson. These 800 times will form the population that you will investigate in this lesson.

Look at the values in the population. Can you find the longest time spent in the gym in the population? Can you find the shortest?

On average, roughly how long do you think people spend at the gym? In other words, by just looking at the numbers in the two tables, make an estimate of the population mean.

You could find the population mean by typing all 800 numbers into a calculator or a computer, adding them up, and dividing by 800. This would be extremely time-consuming, and usually it is not possible to measure every value in a population.

Instead of doing a calculation using every value in the population, we will use a random sample to find the mean of the sample. The sample mean will then be used as an estimate of the population mean.

Example 2: Selecting a Sample Using a Table of Random Digits

The table of random digits provided with this lesson will be used to select items from a population to produce a random sample from the population. The list of digits are determined by a computer program that simulates a random selection of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, or 9. Imagine that each of these digits is written on a slip of paper and placed in a bag. After thoroughly mixing the bag, one slip is drawn and its digit is recorded in this list of random digits. The slip is then returned to the bag and another slip is selected. The digit on this slip is recorded, and then returned to the bag. The process is repeated over and over. The resulting list of digits is called a random-number table.

How could you use a table of random digits to take a random sample?

Step 1: Place the table of random digits in front of you. Without looking at the page, place the eraser end of your pencil somewhere on the table. Start using the table of random digits at the number closest to where your eraser touched the paper. This digit and the following two specify which observation from the population tables will be the first observation in your sample.

For example, suppose the eraser end of your pencil lands on the twelfth number in row 3 of the random digit table. This number is 5 and the two following numbers are 1 and 4. This means that the first observation in your sample is observation number 514 from the population. Find observation number 514 in the population table. Do this by going to Row 51 and moving across to the column heading "4." This observation is 53, so the first observation in your sample is 53.

If the number from the random-number table is any number 800 or greater, you will ignore this number and use the next three digits in the table.

Step 2: Continue using the table of random digits from the point you reached, and select the other four observations in your sample like you did above.

For example, continuing on from the position in the example given in Step 1:

- The next number from the random-digit table is 716, and observation 716 is 63
- The next number from the random-digit table is 565, and observation 565 is 31.
- The next number from the random-digit table is 911, and there is no observation 911. So we ignore these three digits.
- The next number from the random-digit table is 928, and there is no observation 928. So we ignore these three digits.
- The next number from the random-digit table is 303, and observation 303 is 70.
- The next number from the random-digit table is 677, and observation 677 is 42.

Exercises 1–4

Initially, you will select just five values from the population to form your sample. This is a very small sample size, but it is a good place to start to understand the ideas of this lesson.

The values found in Example 2 are used to illustrate the answers to the following questions, but each student should select their own random sample of five.

1. Use the table of random number to select five values from the population of times. What are the five observations in your sample?

2. For the sample that you selected, calculate the sample mean.

3. You selected a random sample and calculated the sample mean in order to estimate the population mean. Do you think that the mean of these five observations is exactly correct for the population mean? Could the population mean be greater than the number you calculated? Could the population mean be less than the number you calculated?

4. In practice, you only take one sample in order to estimate a population characteristic. But, for the purposes of this lesson, suppose you were to take another random sample from the same population of times at the gym. Could the new sample mean be closer to the population mean than the mean of these five observations? Could it be further from the population mean?

Exercises 5–7

As a class, you will now investigate sampling variability by taking several samples from the same population. Each sample will have a different sample mean. This variation provides an example of sampling variability.

5. Place the table of random digits in front of you and, without looking at the page, place the eraser end of your pencil somewhere on the table of random numbers. Start using the table of random digits at the number closest to where your eraser touches the paper. This digit and the following two specify which observation from the population tables will be the first observation in your sample. Write this three-digit number and the corresponding data value from the population in the space below.

6. Continue moving to the right in the table of random digits from the place ended in Exercise 5. Use three digits at a time. Each set of three digits specifies which observation in the population is the next number in your sample. Continue until you have four more observations, and write these four values in the space below.

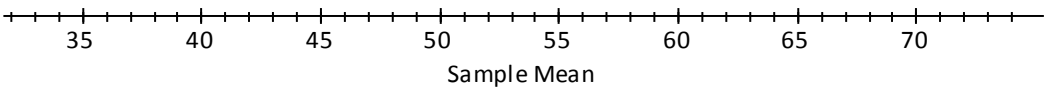
7. Calculate the mean of the five values that form your sample. Round your answer to the nearest tenth. Show your work and your sample mean in the space below.

Exercises 8–11

You will now use the sample means from Exercise 7 from the entire class to make a dot plot.

8. Write the sample means for everyone in the class in the space below.

9. Use all the sample means to make a dot plot using the axis given below. (Remember, if you have repeated or close values, stack the dots one above the other.)



10. What do you see in the dot plot that demonstrates sampling variability?

11. Remember that in practice you only take one sample. (In this lesson, many samples were taken in order to demonstrate the concept of sampling variability.) Suppose that a statistician plans to take a random sample of size 5 from the population of times spent at the gym and that he or she will use the sample mean as an estimate of the population mean. Approximately how far can the statistician expect the sample mean to be from the population mean?

Population

	0	1	2	3	4	5	6	7	8	9
00	45	58	49	78	59	36	52	39	70	51
01	50	45	45	66	71	55	65	33	60	51
02	53	83	40	51	83	57	75	38	43	77
03	49	49	81	57	42	36	22	66	68	52
04	60	67	43	60	55	63	56	44	50	58
05	64	41	67	73	55	69	63	46	50	65
06	54	58	53	55	51	74	53	55	64	16
07	28	48	62	24	82	51	64	45	41	47
08	70	50	38	16	39	83	62	50	37	58
09	79	62	45	48	42	51	67	68	56	78
10	61	56	71	55	57	77	48	65	61	62
11	65	40	56	47	44	51	38	68	64	40
12	53	22	73	62	82	78	84	50	43	43
13	81	42	72	49	55	65	41	92	50	60
14	56	44	40	70	52	47	30	9	58	53
15	84	64	64	34	37	69	57	75	62	67
16	45	58	49	78	59	36	52	39	70	51
17	50	45	45	66	71	55	65	33	60	51
18	53	83	40	51	83	57	75	38	43	77
19	49	49	81	57	42	36	22	66	68	52
20	60	67	43	60	55	63	56	44	50	58
21	64	41	67	73	55	69	63	46	50	65
22	54	58	53	55	51	74	53	55	64	16
23	28	48	62	24	82	51	64	45	41	47
24	70	50	38	16	39	83	62	50	37	58
25	79	62	45	48	42	51	67	68	56	78
26	61	56	71	55	57	77	48	65	61	62
27	65	40	56	47	44	51	38	68	64	40
28	53	22	73	62	82	78	84	50	43	43
29	81	42	72	49	55	65	41	92	50	60
30	56	44	40	70	52	47	30	9	58	53
31	84	64	64	34	37	69	57	75	62	67
32	45	58	49	78	59	36	52	39	70	51
33	50	45	45	66	71	55	65	33	60	51
34	53	83	40	51	83	57	75	38	43	77
35	49	49	81	57	42	36	22	66	68	52
36	60	67	43	60	55	63	56	44	50	58
37	64	41	67	73	55	69	63	46	50	65
38	54	58	53	55	51	74	53	55	64	16
39	28	48	62	24	82	51	64	45	41	47

Population, continued

	0	1	2	3	4	5	6	7	8	9
40	53	70	59	62	33	31	74	44	46	68
41	37	51	84	47	46	33	53	54	70	74
42	35	45	48	45	56	60	66	60	65	57
43	42	81	67	64	60	79	46	48	67	56
44	41	21	41	58	48	38	50	53	73	38
45	35	28	43	43	55	39	75	45	68	36
46	64	31	31	40	84	79	47	63	48	46
47	34	36	54	61	33	16	50	60	52	55
48	53	52	48	47	77	37	66	51	61	64
49	40	44	45	22	36	64	50	49	64	39
50	45	69	67	33	55	61	62	38	51	43
51	55	39	46	56	53	50	44	42	40	60
52	11	36	56	69	72	73	71	48	58	52
53	81	47	36	54	81	59	50	42	80	69
54	40	43	30	54	61	13	73	65	52	40
55	71	78	71	61	54	79	63	47	49	73
56	53	70	59	62	33	31	74	44	46	68
57	37	51	84	47	46	33	53	54	70	74
58	35	45	48	45	56	60	66	60	65	57
59	42	81	67	64	60	79	46	48	67	56
60	41	21	41	58	48	38	50	53	73	38
61	35	28	43	43	55	39	75	45	68	36
62	64	31	31	40	84	79	47	63	48	46
63	34	36	54	61	33	16	50	60	52	55
64	53	52	48	47	77	37	66	51	61	64
65	40	44	45	22	36	64	50	49	64	39
66	45	69	67	33	55	61	62	38	51	43
67	55	39	46	56	53	50	44	42	40	60
68	11	36	56	69	72	73	71	48	58	52
69	81	47	36	54	81	59	50	42	80	69
70	40	43	30	54	61	13	73	65	52	40
71	71	78	71	61	54	79	63	47	49	73
72	53	70	59	62	33	31	74	44	46	68
73	37	51	84	47	46	33	53	54	70	74
74	35	45	48	45	56	60	66	60	65	57
75	42	81	67	64	60	79	46	48	67	56
76	41	21	41	58	48	38	50	53	73	38
77	35	28	43	43	55	39	75	45	68	36
78	64	31	31	40	84	79	47	63	48	46
79	34	36	54	61	33	16	50	60	52	55

Table of Random Digits

Row																				
1	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
2	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
3	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
4	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
5	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
6	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
7	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
8	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7
9	3	4	2	2	9	1	9	0	7	8	1	6	2	5	3	9	0	9	1	0
10	2	7	3	9	5	9	9	3	2	9	3	9	1	9	0	5	5	1	4	2
11	0	2	5	4	0	8	1	7	0	7	1	3	0	4	3	0	6	4	4	4
12	8	6	0	5	4	8	8	2	7	7	0	1	0	1	7	1	3	5	3	4
13	4	2	6	4	5	2	4	2	6	1	7	5	6	6	4	0	8	4	1	2
14	4	4	9	8	7	3	4	3	8	2	9	1	5	3	5	9	8	9	2	9
15	6	4	8	0	0	0	4	2	3	8	1	8	4	0	9	5	0	9	0	4
16	3	2	3	8	4	8	8	6	2	9	1	0	1	9	9	3	0	7	3	5
17	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
18	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
19	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
20	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
21	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
22	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
23	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
24	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7
25	3	4	2	2	9	1	9	0	7	8	1	6	2	5	3	9	0	9	1	0
26	2	7	3	9	5	9	9	3	2	9	3	9	1	9	0	5	5	1	4	2
27	0	2	5	4	0	8	1	7	0	7	1	3	0	4	3	0	6	4	4	4
28	8	6	0	5	4	8	8	2	7	7	0	1	0	1	7	1	3	5	3	4
29	4	2	6	4	5	2	4	2	6	1	7	5	6	6	4	0	8	4	1	2
30	4	4	9	8	7	3	4	3	8	2	9	1	5	3	5	9	8	9	2	9
31	6	4	8	0	0	0	4	2	3	8	1	8	4	0	9	5	0	9	0	4
32	3	2	3	8	4	8	8	6	2	9	1	0	1	9	9	3	0	7	3	5
33	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
34	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
35	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
36	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
37	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
38	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
39	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
40	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7

Lesson Summary

A population characteristic is estimated by taking a random sample from the population and calculating the value of a statistic for the sample. For example, a population mean is estimated by selecting a random sample from the population and calculating the sample mean.

The value of the sample statistic (e.g., the sample mean) will vary based on the random sample that is selected. This variation from sample to sample in the values of the sample statistic is called **sampling variability**.

Problem Set

1. Yousef intends to buy a car. He wishes to estimate the mean fuel efficiency (in miles per gallon) of all cars available at this time. Yousef selects a random sample of 10 cars and looks up their fuel efficiencies on the Internet. The results are shown below.

   22 25 29 23 31 29 28 22 23 27

   a. Yousef will estimate the mean fuel efficiency of all cars by calculating the mean for his sample. Calculate the sample mean, and record your answer below. (Be sure to show your work.)

   b. In practice, you only take one sample to estimate a population characteristic. However, if Yousef were to take another random sample of 10 cars from the same population, would he likely get the same value for the sample mean?

   c. What if Yousef were to take *many* random samples of 10 cars? Would the entire sample means be the same?

   d. Using this example, explain what sampling variability means.

2. Think about the mean number of siblings (brothers and sisters) for all students at your school.

   a. Approximately what do you think is the value of the mean number of siblings for the population of all students at your school?

   b. How could you find a better estimate of this population mean?

   c. Suppose that you have now selected a random sample of students from your school. You have asked all of the students in your sample how many siblings they have. How will you calculate the sample mean?

   d. If you had taken a different sample, would the sample mean have taken the same value?

   e. There are many different samples of students that you could have selected. These samples produce many different possible sample means. What is the phrase used for this concept?

   f. Does the phrase you gave in part (e) apply only to sample means?