Bookmark this page

Learn By Doing Activity

Scenario: Introductory Statistics Students

Consider the population of all students at a large university taking introductory statistics courses (1,129 students taking statistics for business, social sciences, or natural sciences).

Suppose we are interested in the values of four specific variables for this population: handedness (right-handed or left-handed), sex, SAT Verbal score, and age. If we were unable to determine the values of those variables for the entire population, we may be able to take a random sample from that population, and use the sample summaries as estimates for population summaries. Would the random sample provide unbiased estimates for the population values?

Our data set contains data on the entire population of 1,129 students, which includes students taking introductory statistics who are majoring in the natural and social sciences, as well as for business majors.

The variables are defined as follows:

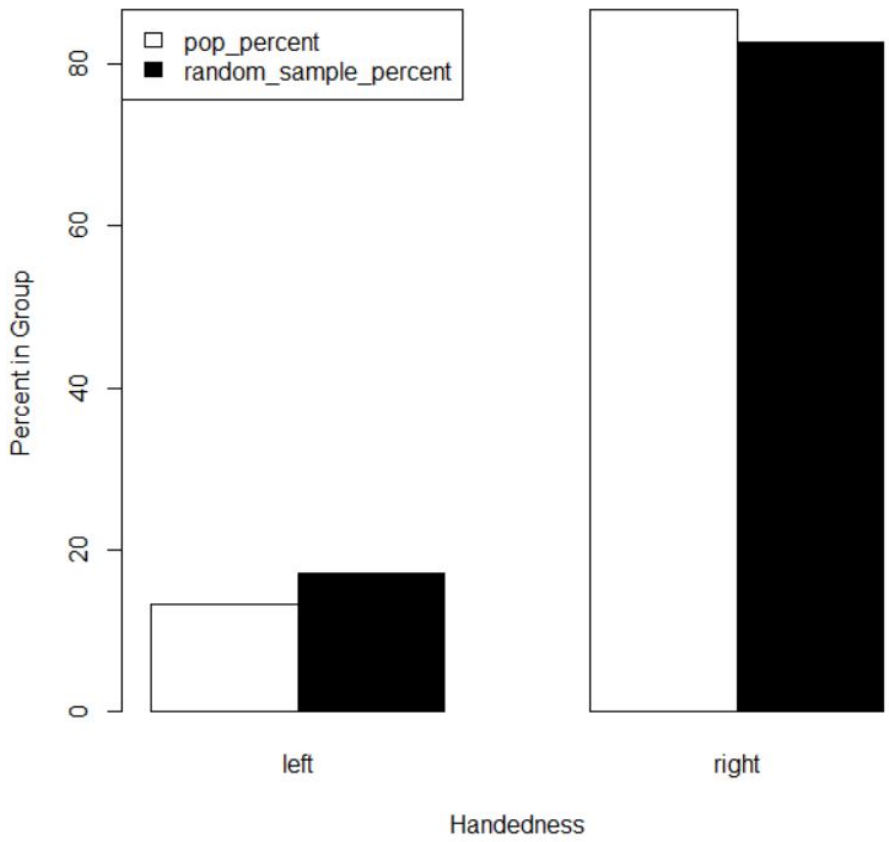
- Course: natural science, social science, or business
- Handed: righthanded or lefthanded
- Sex: female or male
- Verbal: SAT Verbal scores up to 800
- Age: in years

Learn By Doing

1/1 point (graded)

First, we take a simple random sample of the data containing 192 students.

To compare the proportion of right-handed students in the sample to those in the population, we created two side-by-side bar graphs: one for handedness in the population and one for handedness in the random sample.



True or false? The distribution of the sample proportion of handedness is comparable to the population proportion.

- (•) True ✔
- ( ) False

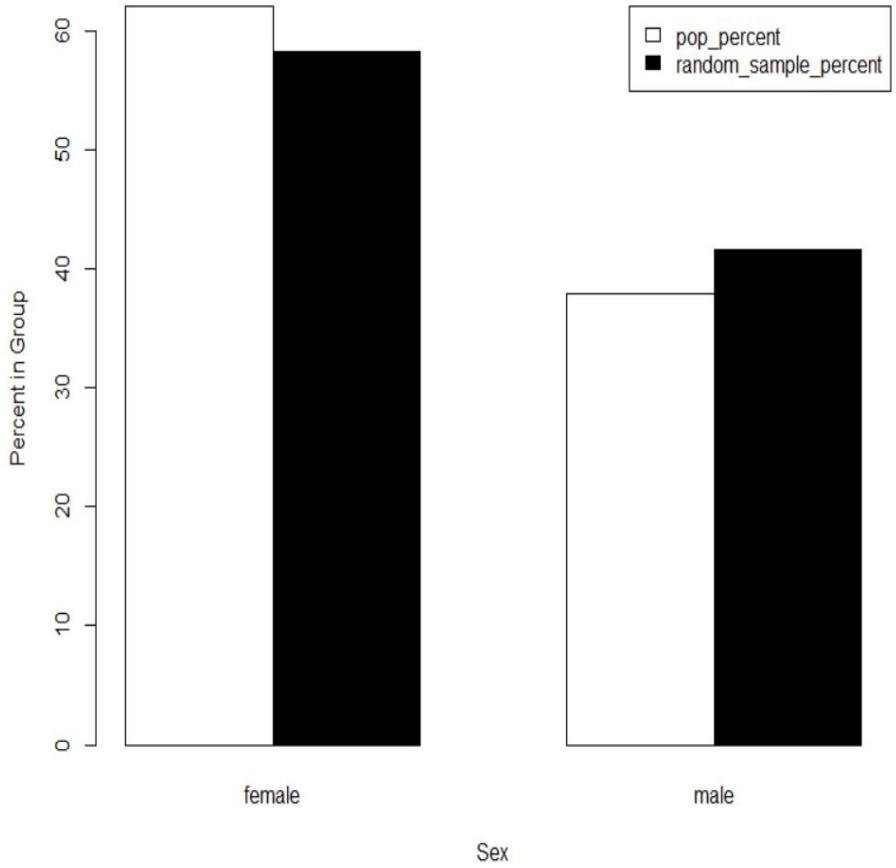
Answer
Correct:
Since the distribution of the sample proportion comes within about 5% of the population proportion they are comparable.

Submit

Learn By Doing

1/1 point (graded)

We used the same random sample of 192 student to compare the proportion of females and males in the sample to those in the population. We created two side-by-side bar graphs: one for sex in the population and one for sex in the random sample.



True or false? The distribution of the sample proportion of sex is NOT comparable to the population proportion.

- True
- False ✔

Answer

Correct:

Since the distributions of the sample proportion comes within about 5% of the population proportion they are comparable.

Submit

We calculated descriptive statistics for SAT Verbal score in the population and for SAT Verbal score in the random sample of 192 students as shown in the table below.

SAT Verbal Scores

Group	Min.	1st Qu.	Median	Mean	3rd Qu.	Max	NA's
Population	300.0	540.0	590.0	588.8	640.0	800.0	131
Random Sample	340.0	520.0	580.0	576.4	630.0	790.0	24

Learn By Doing

1/1 point (graded)

SAT scores tend to follow a normal (symmetric) distribution.

Given the distribution of SAT Verbal scores in the random sample compared to the population, is there evidence of bias in the sample?

- ○ Yes, because the range for the sample (450) is smaller than the range for the population (500).
- ○ Yes, because the IQR for the sample (110) is larger than the IQR for the population (100).
- ◉ No, because the means for the sample and population are approximately equal. ✔
- ○ No, because the proportion of NA's in the sample and population is approximately equal.

Answer

Correct:

Since SAT scores tend to follow a normal distribution, we can use the means to compare the distributions of the sample and population. The means for the two distribution are within 15 points indicating there is NOT evdence of bias in the sample.

Submit

We calculated descriptive statistics for Age in the population and for Age in the random sample of 192 students as shown in the table below.

Age

Group	Min.	1st Qu.	Median	Mean	3rd Qu.	Max	NA's
Population	17.67	19.08	19.67	20.37	20.67	74.17	6

Random Sample	17.92	19.08	19.75	20.59	20.88	74.17	1

Learn By Doing

1/1 point (graded)

Since Age tends to follow a right-skewed distribution, you should focus on medians to make a comparison.

Are the distributions of the sample and population comparable?

- (•) Yes ✔
- ( ) No

Answer

Correct:

Since the sample median age is within 0.5 years of the population median age, the two distributions are comparable.

Submit