Practice Exercises: Lesson 3.1

Diez, D. M., Çetinkaya-Rundel, M., Barr, C. D. (2019). OpenIntro Statistics (4th ed.). OpenIntro. https://www.openintro.org/book/os/ CC BY-SA 3.0

STAT 1201
Introduction to Probability and Statistics

Exercises

5.1 Identify the parameter, Part I. For each of the following situations, state whether the parameter of interest is a mean or a proportion. It may be helpful to examine whether individual responses are numerical or categorical.

(a) In a survey, one hundred college students are asked how many hours per week they spend on the Internet.

(b) In a survey, one hundred college students are asked: "What percentage of the time you spend on the Internet is part of your course work?"

(c) In a survey, one hundred college students are asked whether or not they cited information from Wikipedia in their papers.

(d) In a survey, one hundred college students are asked what percentage of their total weekly spending is on alcoholic beverages.

(e) In a sample of one hundred recent college graduates, it is found that 85 percent expect to get a job within one year of their graduation date.

5.2 Identify the parameter, Part II. For each of the following situations, state whether the parameter of interest is a mean or a proportion.

(a) A poll shows that 64% of Americans personally worry a great deal about federal spending and the budget deficit.

(b) A survey reports that local TV news has shown a 17% increase in revenue within a two year period while newspaper revenues decreased by 6.4% during this time period.

(c) In a survey, high school and college students are asked whether or not they use geolocation services on their smart phones.

(d) In a survey, smart phone users are asked whether or not they use a web-based taxi service.

(e) In a survey, smart phone users are asked how many times they used a web-based taxi service over the last year.

5.3 Quality control. As part of a quality control process for computer chips, an engineer at a factory randomly samples 212 chips during a week of production to test the current rate of chips with severe defects. She finds that 27 of the chips are defective.

(a) What population is under consideration in the data set?

(b) What parameter is being estimated?

(c) What is the point estimate for the parameter?

(d) What is the name of the statistic can we use to measure the uncertainty of the point estimate?

(e) Compute the value from part (d) for this context.

(f) The historical rate of defects is 10%. Should the engineer be surprised by the observed rate of defects during the current week?

(g) Suppose the true population value was found to be 10%. If we use this proportion to recompute the value in part (e) using $p = 0.1$ instead of $\hat{p}$, does the resulting value change much?

5.4 Unexpected expense. In a random sample 765 adults in the United States, 322 say they could not cover a $400 unexpected expense without borrowing money or going into debt.

(a) What population is under consideration in the data set?

(b) What parameter is being estimated?

(c) What is the point estimate for the parameter?

(d) What is the name of the statistic can we use to measure the uncertainty of the point estimate?

(e) Compute the value from part (d) for this context.

(f) A cable news pundit thinks the value is actually 50%. Should she be surprised by the data?

(g) Suppose the true population value was found to be 40%. If we use this proportion to recompute the value in part (e) using $p = 0.4$ instead of $\hat{p}$, does the resulting value change much?

5.5 Repeated water samples. A nonprofit wants to understand the fraction of households that have elevated levels of lead in their drinking water. They expect at least 5% of homes will have elevated levels of lead, but not more than about 30%. They randomly sample 800 homes and work with the owners to retrieve water samples, and they compute the fraction of these homes with elevated lead levels. They repeat this 1,000 times and build a distribution of sample proportions.

(a) What is this distribution called?

(b) Would you expect the shape of this distribution to be symmetric, right skewed, or left skewed? Explain your reasoning.

(c) If the proportions are distributed around 8%, what is the variability of the distribution?

(d) What is the formal name of the value you computed in (c)?

(e) Suppose the researchers' budget is reduced, and they are only able to collect 250 observations per sample, but they can still collect 1,000 samples. They build a new distribution of sample proportions. How will the variability of this new distribution compare to the variability of the distribution when each sample contained 800 observations?

5.6 Repeated student samples. Of all freshman at a large college, 16% made the dean's list in the current year. As part of a class project, students randomly sample 40 students and check if those students made the list. They repeat this 1,000 times and build a distribution of sample proportions.

(a) What is this distribution called?

(b) Would you expect the shape of this distribution to be symmetric, right skewed, or left skewed? Explain your reasoning.

(c) Calculate the variability of this distribution.

(d) What is the formal name of the value you computed in (c)?

(e) Suppose the students decide to sample again, this time collecting 90 students per sample, and they again collect 1,000 samples. They build a new distribution of sample proportions. How will the variability of this new distribution compare to the variability of the distribution when each sample contained 40 observations?