9.3 Exercises

Basic Concepts

1. What is a proportion? What type of information does it give us about the population?
2. How is the sample proportion found?
3. Describe, in layman's terms, how a confidence interval is constructed for a population proportion.
4. It seems that estimating proportions produces estimates which are much more precise than those for means. Explain why this is the case.
5. The population proportion is often unknown. How is this issue dealt with when determining sample size?
6. What is the guideline to follow when there is no estimate available for the population proportion? Why is this done?
7. How do the resulting required sample sizes differ when there is an estimate available versus when there is no estimate available for the population proportion?

Exercises

8. Acid rain accumulations in lakes and streams in the northeastern part of the United States are a major environmental concern. A researcher wants to know what fraction of lakes contain hazardous pollution levels. He randomly selects 200 lakes and determines that 45 of the selected lakes have an unsafe concentration of acid rain pollution.
    a. Calculate the best point estimate of the population proportion of lakes that have unsafe concentrations of acid rain pollution.
    b. Determine a 95% confidence interval for the population proportion.
    c. If a local politician states that only 20% of the lakes are contaminated, does the study provide overwhelming evidence at the 95% level to contradict his views?
9. *The Richland Gazette*, a local newspaper, conducted a poll of 1000 randomly selected readers to determine their views concerning the city's handling of snow removal. The paper found that 650 people in the sample felt the city did a good job.
    a. Compute the best point estimate for the percentage of readers who believe the city is doing a good job of snow removal.
    b. Construct a 90% confidence interval for this percentage.
10. The clinical testing of drugs involves many factors. For example, patients that have been given placebos, which are harmless compounds that have no effect on the patient, often will still report that they feel better. Assume that in a study of 500 random subjects conducted by the Poppins Sucre Drug Company, the percentage of patients reporting improvement when given a placebo was 37%.
    a. What would be a 95% confidence interval for the true proportion of patients who exhibit the placebo effect? Interpret this interval in terms of the problem.
    b. What would the 99% confidence interval be?
    c. To gain the additional 4% of confidence how much wider did the interval become?
11. The Peacock Cable Television Company thinks that 40% of their customers have more outlets wired than they are paying for. A random sample of 400 houses reveals that 110 of the houses have excessive outlets.
    a. Construct a 99% confidence interval for the true proportion of houses having too many outlets.

b. Do you feel the company is accurate in its belief about the proportion of customers who have more outlets wired than they are paying for? Justify your answer.

12. Running continues to be a very popular sport in America. At a major race, like the Peachtree Road Race in Atlanta, there may be over 10,000 people entered to run. The race promoters for a road race in the Pacific Northwest took a random sample of 750 runners out of the 5000 runners entered to estimate the number of runners who will need hotel accommodations. Five hundred runners indicated they would need hotel accommodations.

a. Construct a 90% confidence interval for the true proportion of runners who will need hotel accommodations.

b. Is the confidence interval obtained sufficiently narrow to be of help in planning the number of hotel rooms which will be necessary to accommodate the runners? Justify your answer.

13. In the fourth quarter of 2010 the home ownership rate was 66.5%. This rate is 2.7 percentage points lower than the 2004 peak of 69.2%, and the lowest rate since 1998. Home ownership fell at an alarming pace in the fourth quarter of the year, despite the fact that home prices fell, affordability was much improved, and inventories of new and existing homes were running quite high. Suppose that a random sample of 120 households was selected from an area in the Midwest that is particularly economically depressed. Suppose that 57 of the households sampled were owned by the residents of the homes.
Source: U.S. Census Bureau

a. Construct a 95% confidence interval for the proportion of households in the area sampled that are owned by the residents of the homes.

b. Is there evidence at the 95% level that the proportion of the households in the area sampled that are owned by residents is less than the national rate?

14. In the *Gallup Poll Monthly*, it was reported that 31% of the people surveyed in a recent poll claimed that vegetables were their least favorite food. Surprisingly, only 14% responded with liver, and 10% of those surveyed did not submit a response because they claimed that they liked everything. The poll was based upon a sample of 1001 people. Assuming that a random sample was chosen, construct a 90% confidence interval for the percentage of all Americans who say that vegetables are their least favorite food.

15. The Federal Trade Commission (FTC) conducted a study investigating the accuracy of bar-code scanners. It was concluded that these computer scanners, used mainly at grocery, department store, and drugstore checkout counters, ring up the wrong price about 5% of the time. In most instances, however, the error was in favor of the shopper, according to the FTC. Suppose that your local grocery store conducts a study to determine the accuracy of its scanners. Assume 13 shoppers are randomly chosen and their bills, as indicated by the scanner, are checked against the correct bill computed by conventional means. Suppose that of the 200 items scanned, 21 of the items were charged incorrectly by the scanner.

a. Construct a 95% confidence interval for the proportion of items that were rung up incorrectly by the scanner.

b. Does it appear that the local grocery store has a larger error rate than 5%?

16. The Big Green Poster Company wants to estimate the fraction of poster sites controlled by their competition, Bird's Billboard Service. What sample size would be necessary to estimate this fraction to within 3% with 95% confidence? (They think Bird's controls about 33 percent of the boards.)

17. Researchers working in a remote area of Africa feel that 40% of families in the area are without adequate drinking water either through contamination or unavailability. What sample size will be necessary to estimate the percentage without adequate water to within 5% with 99% confidence?

18. Companies that provide environmental cleanup for hazardous waste and toxic chemicals are growing rapidly. W.R. Gross is thinking about entering this field with a subsidiary called Saf-t-Soil. They wish to estimate the true proportion of U.S. corporations that produce hazardous waste as a by-product of their manufacturing process to within 10% with 80% confidence. What sample size will be needed?

19. The public relations manager for a political candidate would like to determine if the registered voters in the candidate's district agree with the politician's view on a particular issue. Find the sample size necessary for the public relations manager to estimate the true proportion to within 5% with 85% confidence.

9.4 Estimating the Population Standard Deviation or Variance

Recall that the sample variance is

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

and it serves as the point estimate of the population variance, σ^2. As with the other tests developed for the population mean and the population proportion, we first need to develop a sampling distribution for

$$\frac{(n-1)s^2}{\sigma^2}$$

that will allow us to calculate a confidence interval for the population variance.

Formula

χ^2 Test Statistic

If we have a random sample of size n taken from a normal population, then the sampling distribution of the test statistic is given by

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

which has a **chi-square distribution** with $n - 1$ degrees of freedom.

The chi-square distribution is a positively skewed (or skewed to the right) distribution. Like the *t*-distribution, the shape of the distribution is a function of its degrees of freedom. See Figure 9.4.1 which illustrates the chi-square distributions with 4 and 10 degrees of freedom.

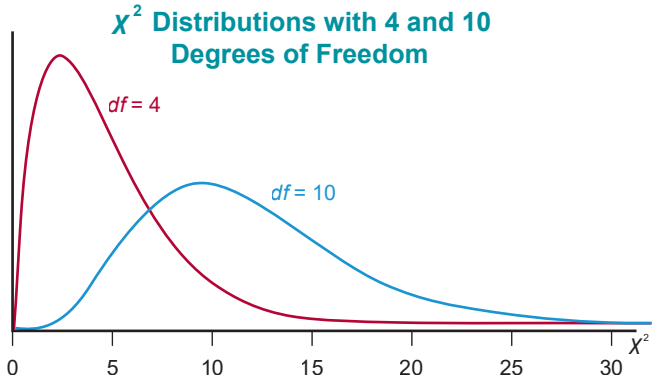

Figure 9.4.1