What Is a Confidence Interval Anyway?

Explaining the confidence interval as the range of plausible population values.

Forty people are selected at random and given a test to identify their dominant eye. The person holds an 8.5 × 11 piece of paper with about a 1 × 1 inch square cut in the middle at arm's length with both hands. The person looks through the square at a relatively small object across the room. The person then closes one eye. If he or she can still see the object, the open eye is the dominant eye. If not, the closed eye is the dominant eye.

Question

Is this sample of only 40 people large enough for us to come to any conclusion about what percentage of people have a dominant right eye?

Objectives

In this activity, you will learn how to construct confidence intervals using simulation. You will also learn how to interpret confidence intervals.

Prerequisites

You should know how to simulate samples from a given binomial population by using a random number table.

Activity

1. Conduct the dominant-eye experiment just described with 40 students from your class, adding other people as necessary to bring the total up to 40. What proportion of your sample was right-eye dominant?

2. In this activity, you will be taking samples from a *population* in which 30% have some characteristic in order to see how close the proportions in the *samples* tend to come to 30%.
   a. Use a random number table to simulate taking a sample of size 40 from a population with 30% "successes." Let the digits 0, 1, and 2 represent a success and 3, 4, 5, 6, 7, 8, and 9 represent a failure. Place a tally mark in a table like the one shown here to represent your result.

Number of Successes	Frequency	Number of Successes	Frequency
0		21	
1		22	
2		23	
3		24	
4		25	
5		26	
6		27	
7		28	
8		29	
9		30	
10		31	
11		32	
12		33	
13		34	
14		35	
15		36	
16		37	
17		38	
18		39	
19		40	
20		Total	100

Table 1

b. Combine your results with other members of your class, repeating the simulation until your class has placed tally marks from 100 different samples in the frequency column of Table 1.

c. Comparing your proportion from step 1 with the frequency table from step 2b, is it plausible (that is, is there a reasonable chance) that a sample of 40 drawn from a population where 30% are right-eye dominant would have the number of right-eyed people we see in our sample? Explain.

d. Complete the following sentences based on your frequency table from step 2b: Less than 5% of the time, there were _____ successes or fewer. Less than 5% of the time, there were _____ successes or more.

e. In Figure 1, draw a thin horizontal box aligned with the 30% "Percentage of Successes in Population" on each side. Using "Number of Successes in Sample" on the bottom as a guide, the box should stretch in length between your two answers to step 2d.

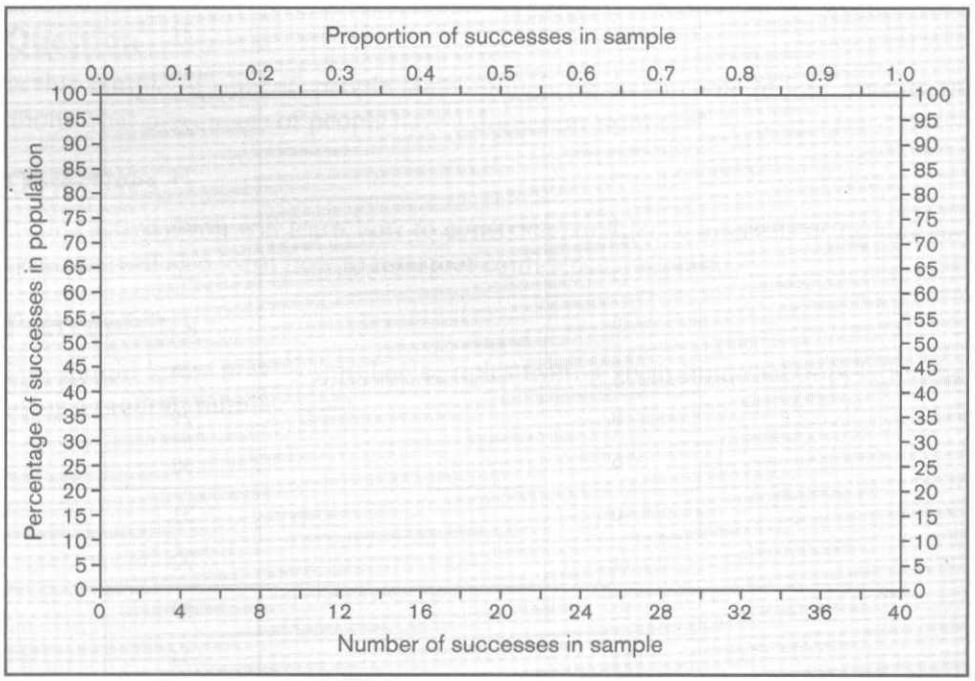

Figure 1

f. About 30% of Americans aged 19 to 28 claim that they have used an illicit drug other than marijuana. If a random sample of 40 Americans aged 19 to 28 finds 21 who claim to have used an illicit drug other than marijuana, would you be surprised? Explain.

g. According to the U.S. Bureau of Labor Statistics, about 30% of women with children younger than 6 years of age do not participate in the labor force. Would it be plausible for a survey of 40 randomly chosen mothers of children younger than 6 to find that 8 are not working? Explain.

3. The class should now divide into groups of about five students each. Your instructor will give each group one of the percentages on the sides of Figure 1. With your group, repeat steps 2a, 2b, 2d, and 2e for your new percentage.

4. Use the completed chart to answer these questions.
   a. According to the 1990 U.S. Census, about 30% of people aged 25 to 44 live alone. In a random sample of 40 people aged 25 to 44, would it be plausible to get 20 who live alone?
   b. We select a random sample of size 40 and get a sample proportion of 0.90 successes. Is this sample proportion plausible if the population has 75% successes?
   c. According to the 1990 U.S. Census, about 80% of men aged 20 to 24 have never been married. In a random sample of 40 men aged 20 to 24, how many unmarried men is it plausible to get? What proportion of unmarried men is it plausible to get?
   d. Suppose you flip a coin 40 times; how many heads is it plausible for you to get?
   e. In the 1992 presidential election, Bill Clinton got 43% of the vote. In a random sample of 40 voters, what is the largest proportion of people who voted for Clinton that it is plausible to get? The smallest?

5. Based on your sample and on the chart your class has completed, is it plausible that 60% is the percentage of the population is right-eye dominant? What percentages are plausible? These percentages are called the *90% confidence interval.*

6. Use your completed chart to find these confidence intervals.
   a. Suppose that a random sample of 40 toddlers finds that 34 know what color Barney is. What is the 90% confidence interval for the percentage of toddlers who know what color Barney is?
   b. Suppose that a random sample of 40 adults finds that 10 know what color Barney is. What is the 90% confidence interval for the percentage of adults who know what color Barney is?
   c. Observe 40 students on your campus. Find the 90% confidence interval for the percentage of students who carry backpacks.

Wrap-Up

1. Polls usually report a margin of error. Suppose a poll of 40 randomly selected statistics majors finds that 20 are female. The poll reports that 50% of statistics majors are female, with a margin of error of 10%. Use your completed chart to explain where the 10% came from.

2. People often complain that election polls cannot be right because they personally were not asked how they were going to vote. Write an explanation to such a person about how polls can get a good idea of how the entire population will vote by asking a relatively small number of voters.

Extensions

1. This activity used 90% confidence intervals because it is easy computationally to find the bottom 5% and the top 5% of a distribution. Usually, 95% confidence intervals are reported. Will 95% confidence intervals be longer or shorter than 90% confidence intervals? Explain.

2. Will the confidence intervals for samples of size 80 be longer or shorter than those for samples of size 40? Design and carry out the simulation needed to answer this question.

Assessment Questions

1. A random sample of 40 statistics students at a large university finds that 27 think the textbook is inscrutable. If you were to ask all statistics students at that university if they think their textbook is inscrutable, explain what you would expect to find.

2. You poll a random sample of 40 people and find that none of them have been drivers in car crashes that totaled the car. A fellow statistics student says, "Strange, since the sample proportion is zero, that means that the confidence interval—which gets narrow near 0 and 1—has zero width. It extends from 0 to 0." Use your understanding of confidence intervals from this activity to explain why your fellow student is wrong—and find the 90% confidence interval for this situation.

3. The most difficult idea you will probably encounter in an introductory statistics course is that of a confidence interval. What makes this idea difficult to understand?