

1. The Harris Poll asked a sample of smokers, "Do you believe that smoking will probably shorten your life, or not?" Of the 1010 people in the sample, 848 said "Yes."

(a) Harris called residential telephone numbers at random in an attempt to contact an SRS of smokers. Based on what you know about national sample surveys, what is likely to be the biggest weakness in the survey?

(b) We will nonetheless act as if the people interviewed are an SRS of smokers. Give a 95% confidence interval for the percent of smokers who agree that smoking will probably shorten their lives.

2. Students are reluctant to report cheating by other students. A student project put this question to an SRS of 172 undergraduates at a large university: "You witness two students cheating on a quiz. Do you go to the professor?" Only 19 answered "Yes." Give a 95% confidence interval for the proportion of all undergraduates at this university who would report cheating.

3. In a Harris Poll survey of smokers, 848 of a sample of 1010 smokers agreed that smoking would probably shorten their lives. Harris announces a margin of error of ± 3 percentage points for all samples of about this size. Opinion polls announce the margin of error for 95% confidence.

(a) What is the actual margin of error (in percent) for the large-sample confidence interval from this sample?

(b) The margin of error is largest when $\hat{p} = 0.5$. What would the margin of error (in percent) be if the sample had resulted in $\hat{p} = 0.5$?

(c) Why do you think that Harris announces a $\pm 3\%$ margin of error for all samples of about this size?

4. Plain type fonts such as Times New Roman are easier to read than fancy fonts such as Gigi. A group of 25 volunteer subjects read the same text in both fonts. (This is a matched pairs design. One-sample procedures for proportions, like those for means, are used to analyze data from matched pairs designs.) Of the 25 subjects, 17 said that they preferred Times New Roman for Web use. But 20 said that Gigi was more attractive.

(a) Because the subjects were volunteers, conclusions from this sample can be challenged. Show that the sample size condition for the large-sample confidence interval is not met, but that the condition for the plus four interval is met.

(b) Give a 95% confidence interval for the proportion of all adults who prefer Times New Roman for Web use. Give a 90% confidence interval for the proportion of all adults who think Gigi is more attractive.

5. The Internal Revenue Service plans to examine an SRS of individual federal income tax returns from each state. One variable of interest is the proportion of returns claiming itemized deductions. The total number of tax returns in a state varies from more than 15 million in California to fewer than 250,000 in Wyoming.

(a) Will the margin of error for estimating the population proportion change from state to state if an SRS of 2000 tax returns is selected in each state? Explain your answer.

(b) Will the margin of error change from state to state if an SRS of 1% of all tax returns is selected in each state? Explain your answer.

6. An automobile manufacturer would like to know what proportion of its customers are not satisfied with the service provided by the local dealer. The customer relations department will survey a random sample of customers and compute a 99% confidence interval for the proportion who are not satisfied.

(a) Past studies suggest that this proportion will be about 0.2. Find the sample size needed if the margin of error of the confidence interval is to be about 0.015.

(b) When the sample of the size found in (a) is actually contacted, 10% of the sample say they are not satisfied. What is the margin of error of the 99% confidence interval?

7. The National Assessment of Educational Progress (NAEP) includes a "long-term trend" study that tracks reading and mathematics skills over time and obtains demographic information. In the 2008 study, a random sample of 9600 17-year-old students was selected. The NAEP sample used a multistage design, but the overall effect is quite similar to an SRS of 17-year-olds who are still in school.

(a) In the sample, 46% of students had at least one parent who was a college graduate. Estimate with 99% confidence the proportion of all 17-year-old students in 2008 who had at least one parent who was a college graduate.

(b) The sample does not include 17-year-olds who dropped out of school, so your estimate is valid only for students. Do you think that the proportion of all 17-year-olds with at least one parent who was a college graduate would be higher or lower than 46%? Explain.

8. A sample survey funded by the National Science Foundation asked a random sample of American adults about biological evolution. One question asked subjects to answer "True," "False," or "Not sure" to the statement "Human beings, as we know them today, developed from earlier species of animals." Of the 1484 respondents, 594 said "True." What can you say with 95% confidence about the percent of all American adults who think that humans developed from earlier species of animals?

9. Does the order in which wine is presented make a difference? Several choices of wine are presented one at a time, and the subject is then asked to choose his or her preferred wine at the end of the sequence. In one study, subjects were asked to taste two wine samples in sequence. Both samples given to a subject were the same wine, although the subjects were expecting to taste two different samples of a particular variety. Of the 32 subjects in the study, 22 selected the wine presented first when presented with two identical wine samples.

Do the data give good reason to conclude that the subjects are not equally likely to choose either of the two wines when presented with two identical wine samples in sequence? Can we generalize our conclusions to all wine tasters? Explain.

10. Does the order in which wine is presented make a difference? Several choices of wine are presented one at a time, and the subject is then asked to choose his or her preferred wine at the end of the sequence. In one study, subjects were asked to taste two wine samples in sequence. Both samples given to a subject were the same wine, although the subjects were expecting to taste two different samples of a particular variety. Of the 32 subjects in the study, 22 selected the wine presented first when presented with two identical wine samples.

Suppose now that you would like to duplicate this study with a larger sample size. How large a sample would be needed to obtain margin of error ± 0.05 in the study of choice order for tasting wine? Use the \hat{p} from the above as your guess for the unknown p .