

Practice OE - confidence intervals

1. A survey will be conducted to examine the educational level of adult heads of households in the U.S. Each respondent in the survey will be placed into one of the following two categories:

- Does not have a high school diploma
- Has a high school diploma

The survey will be conducted using a telephone interview. Random-digit dialing will be used to select the sample.

A pilot survey indicated that about 22 percent of the population of adult heads of households do not have a high school diploma. Using this information, how many respondents should be obtained if the goal of the survey is to estimate the proportion of the population who do not have a high school diploma to within 0.03 with 95 percent confidence?

Handwritten calculations for problem 1:

- Margin of error: $\pm .03$
- Confidence interval: $.22 \pm .03$
- Formula: $1.96 \sqrt{\frac{.22(1-.22)}{n}} = .03$
- Result: $n = 734$ respondents
- Intermediate calculation: $\frac{.22(1-.22)}{n} = .000234$
- Final result: $n = 733.333$

2. Researchers at a large health maintenance organization (HMO) are planning a study of a certain mild illness. They will select a random sample of 2000 patients who are ages 35 to 54 and see if they contract the illness in the next year. The researchers are interested in estimating the proportions of men and of women who are likely to develop the illness in each of 4 age-groups. The researchers find the following numbers within each gender and age-group.

	age group			
	35-39	40-44	45-49	50-54
Male	350	230	150	60
Female	445	370	245	150

a) Suppose that at the end of the study, 10 percent of the females in the 40-44 age-group contracted the illness. Calculate a 95 percent confidence interval to estimate the population proportion of females in this age-group that contracted the illness.

Handwritten calculations for problem 2a:

- Sample size: $n = 370$
- Sample proportion: $\hat{p} = .10$
- Confidence interval: $.10 \pm 1.96 \sqrt{\frac{.10(1-.10)}{370}}$
- Result: $(.0694, .1306)$
- Conditions check:
 - State and verify (check) conditions
 - $n \cdot p \geq 10$ and $n(1-p) \geq 10$
 - $370(.10) \geq 10$ $370(1-.10) \geq 10$
 - $37 \geq 10$ $333 \geq 10$
 - SRS from pop. of interest
 - problem states a random sample of 2,000 from all HMO patients (cond. met)

Practice OE - confidence intervals

- b) Interpret this confidence interval in the context of this situation.

I am 95% confident that the actual prop of all of the 40-44 yr. old females from this HMO that would contract the disease is between .07 and .13. \downarrow Variable

population \rightarrow

- c) Interpret the confidence level of 95 percent.

If I take repeated samples of size 370 from this population and create 95% C.I. for each, in the long run, approx. 95% of these intervals would capture the actual prop of this HMO's females 40-44 that would contract the disease.