confidence interval for a proportion

Dr. Mine Çetinkaya-Rundel
Duke University

from the 2010 GSS

Two scientists want to know if a certain drug is effective against high blood pressure. The first scientist wants to give the drug to 1000 people with high blood pressure and see how many of them experience lower blood pressure levels. The second scientist wants to give the drug to 500 people with high blood pressure, and not give the drug to another 500 people with high blood pressure, and see how many in both groups experience lower blood pressure levels. Which is the better way to test this drug?

(a) All 1000 get the drug
(b) 500 get the drug, 500 don't

experimental design	
bad intuition	99
good intuition	571
total	670

What percent of Americans have good intuition about experimental design?

parameter of interest

Percentage of **all** Americans who have good intuition about experimental design.

p

point estimate

Percentage of **sampled** Americans who have good intuition about experimental design.

$\hat{p}$

$571 / 670 \approx 0.85$

estimating a proportion

point estimate ± margin of error

$$\hat{p} \pm z^{\star} SE_{\hat{p}}$$

Standard error for a proportion, for calculating a confidence interval: $SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

practice

The GSS found that 571 out of 670 (~85%) of Americans answered the question on experimental design correctly. Estimate (using a 95% confidence interval) the proportion of all Americans who have good intuition about experimental design?

1. independence: 670 < 10% of Americans, and GSS samples randomly

Whether one American in the sample has good intuition about experimental design is independent of another.

2. sample size / skew: 571 successes, 670 - 571 = 99 failures

Since the success-failure condition is met, we can assume that the sampling distribution of the proportion is nearly normal.

practice

$$\hat{p} \pm z^* \, SE = 0.85 \pm 1.96 \sqrt{\frac{0.85 \times 0.15}{670}}$$

$$= 0.85 \pm 1.96 \times 0.0138$$

$$= 0.85 \pm 0.027$$

$$= (0.823, 0.877)$$

We are 95% confident that 82.3% to 87.7% of all Americans have good intuition about experimental design.

practice

The margin of error for the previous confidence interval was 2.7%. If, for a new confidence interval based on a new sample, we wanted to reduce the margin of error to 1% while keeping the confidence level the same, at least how many respondents should we sample?

$$ME = 0.01 = 1.96\sqrt{\frac{0.85 \times 0.15}{n}}$$

$$0.01^2 = \frac{1.96^2 \times 0.85 \times 0.15}{n}$$

$$n = \frac{1.96^2 \times 0.85 \times 0.15}{0.01^2} = 4898.04 \rightarrow \text{at least } 4899$$

calculating the required sample size for desired ME

remember $ME = z^\star \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

- if there is a previous study that we can rely on for the value of $\hat{p}$ use that in the calculation of the required sample size
- if not, use $\hat{p} = 0.5$
  - if you don't know any better, 50-50 is a good guess
  - gives the most conservative estimate – highest possible sample size