Module 5 Solutions

Xihan Qian

June. 4th, 2022

1. In a small town, some people are complaining that the park is too small and should be expanded. To gather more information about this issue, the managers of the park hand out questionnaires to people who brought their children to the park and ask for their opinions. Discuss possible biases in this sample. [variations of *Stats: Data and Models*: Ch. 10]

Solution: They will only get responses from those people who come to the park, who mostly like are satisfied with the park and won't get any opinions from those who are dissatisfied with the park. They will also only get opinions from those people with children.

2 A random sample of 1000 teenagers were interviewed about their average daily phone use. About 60% said that they would spent around 5-7 hours on their phones per day. [variations of *Stats: Data and Models*: Ch. 16]
(a) Explain what $\hat{p}$ means in this context.
(b) Calculate the standard error of $\hat{p}$.
(c) Explain what the standard error means in this context.

Solution: (a) $\hat{p} = 0.6$ in this context means that 60% of the 1000 teenagers in our sample said they would spend 5-7 hours on their phones on average per day. This is an estimate of p, which is the true proportion of all U.S. teenagers who would say that they do so.
(b) The standard error can be calculated as below:

$$SE(\hat{p}) = \sqrt{\frac{0.6 \times 0.4}{1000}} \approx 0.015$$

(c) The standard error is the estimate of the amount of the variation in the sample proportion we expect to see from sample to sample when we ask 1000 teenagers whether they spend 5-7 hours on average on their phones daily.

3. According to a survey, around 38% of American teenagers are scared to ride the roller coaster. This survey was conducted based on a random sample of 800 teenagers. Is it reasonable for the researchers to use a normal model for the distribution of the sample proportion? Explain. [variations of *Stats: Data and Models*: Ch. 16]

Solution: Since this sample is random, we have the condition satisfied for independence. Now we check $n\hat{p} = 800 \times 0.38 = 304 \geq 10$; $n(1-\hat{p}) = 800 \times 0.62 = 496 \geq 10$. Therefore the success/failure conditions are also met. Then it is safe for us to apply CLT and use a normal model for the distribution of the sample proportion.

4. For a survey asking about teenagers' daily average phone use, a 95% confidence interval was constructed to be [57%, 63%]. Interpret the interval in this context. [variations of *Stats: Data and Models*: Ch. 16]

Solution: If we were to repeat this procedure many times, then about 95% of the confidence intervals we construct would include the true proportion p. / We are 95% confident that the true proportion p lies in 57% and 63%.

5. A delivery company is interested in finding out whether their deliveries are actually arriving on time. They gave calls to a random sample of customers and constructed a 95% confidence interval, which showed that the proportion of orders arriving on time is [80%, 86%]. Interpret whether the following statements are true. Explain why.

(a) Between 80% and 86% deliveries arrive on time.
(b) There is a 95% probability that the true proportion would lie in between 80% and 86%.
(c) If we repeatedly take a random sample and construct a CI with this confidence procedure, 95% of the intervals would contain the true proportion.
(d) 95% of all random samples of customers will show that 83% of the deliveries arrive on time.

Solution: (a) False. This interpretation would imply certainty, which is not correct.
(b) False. Once a confidence interval is constructed, there is no randomness anymore, so we cannot talk about probability/chance.
(c) This is the correct interpretation of a 95% confidence interval.
(d) False. Different samples will give different confidence intervals; and 95% is a proportion of all intervals constructed in this confidence procedure, but not of the random samples.

6. In constructing a confidence interval, the margin of error (MOE) is defined as $z^* \times SE$. Suppose a poll result states that the approval rate for student council is 58%. If we want an MOE of 5% for the next poll with 95% confidence, how many people should be sampled?

Solution: Since we want it with 95% confidence, $z^* = 1.96$. Hence

$$0.05 = z^* \times SE = 1.96 \times \sqrt{\frac{0.58 \times 0.42}{n}}$$

Solving for n, we obtain that

$$n = \frac{1.96^2(0.58)(0.42)}{0.05^2} = 374.33$$

Hence we would need to sample 375 people.

7 A research firm wants to find the proportion of adults who prefer traditional books to E-books. They took a random sample of 200 adults and asked for their opinions. The results show that 65% of adults like E-books better.
(a) What is the standard error?

(b) Check the conditions for constructing an 80% confidence interval.

(c) If you were to construct an 80% interval, what is the value of z^*?

Solution: (a) $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.65\times0.35}{200}} \approx 0.034$

(b) Since this is a random sample, the independence condition is met. For the S/F condition: $n\hat{p} = 200\times0.65 = 130 \geq 10$; $n(1-\hat{p}) = 200 \times 0.35 = 70 \geq 10$. Hence this condition is also met.