Module 7 Solutions

Xihan Qian

June, 2022

1. An organization is interested in whether people were born in the country that they are currently living in. They randomly surveyed 980 U.S. residents and found out that 196 of them were born in another country. Similarly, they also interviewed 1560 randomly chosen Chinese residents while getting the result that 212 were not born in China.

(a) Compute the standard error of the difference in sample proportions.

Solutions:

$$\hat{p}_1 = \frac{196}{980} = 0.2; \hat{p}_2 = \frac{212}{1560} = 0.136$$

$$\text{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}} = \sqrt{\frac{0.2 \times 0.8}{980} + \frac{0.136 \times 0.864}{1560}} = 0.0154$$

(b) Check the conditions that need to be satisfied in order to compute confidence intervals.

Solutions: Independence within/across groups is met because the samples are both random.

$$\text{S/F condition: } n_1\hat{p}_1 = 196 \geq 10; n_1(1-\hat{p}_1) = 784 \geq 10; n_2\hat{p}_2 = 212 \geq 10; n_2(1-\hat{p}_2) = 1348 \geq 10$$

(c) Construct a 95% two proportion confidence interval for the difference between U.S. and Chinese residents who were born in a foreign country.

Solutions:

$$I = (\hat{p}_1 - \hat{p}_2 - z^* \times SE, \hat{p}_1 - \hat{p}_2 + z^* \times SE) = (0.2 - 0.136 - 1.96 \times 0.0154, 0.2 - 0.136 + 1.96 \times 0.0154) = (3.38\%, 9.42\%)$$

(d) Interpret the confidence interval computed in context.

Solutions: We are 95% confidence that based on these data, the proportion of foreign-born American residents is between 3.38% and 9.42% higher than the proportion of foreign-born Chinese residents.

2. Continuing in the setting from Exercise 1, we want to test whether the proportion of people that were born in another country is the same in the U.S. and China.

Solutions: (a) Let p_1 denote the proportion of Americans that are born outside of the U.S. and p_2 be the proportion of Chinese that are born outside of China.

H_0 : proportion of people that were born in another country is the same for the U.S. and China

$$p_1 - p_2 = 0;$$

H_A : proportion of people that were born in another country is different for the U.S. and China

$$p_1 - p_2 \neq 0;$$

(b)

$$\hat{p}_{pooled} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{196 + 212}{980 + 1560} = 0.1606$$

$$SE = \sqrt{\frac{\hat{p}_{pooled}\left(1 - \hat{p}_{pooled}\right)}{n_1} + \frac{\hat{p}_{pooled}\left(1 - \hat{p}_{pooled}\right)}{n_2}} = \sqrt{\frac{0.1606 \times 0.8394}{980} + \frac{0.1606 \times 0.8394}{1560}} = 0.0150$$

(c) Independence within/across groups is met because the samples are both random.

$$\text{S/F condition: } n_1\hat{p}_{pooled} = 980 \times 0.1606 = 157.39 \geq 10; n_1(1 - \hat{p}_{pooled}) = 822.61 \geq 10;$$

$$n_2\hat{p}_{pooled} = 1560 \times 0.1606 = 250.54 \geq 10; n_2(1 - \hat{p}_{pooled}) = 1309.46 \geq 10$$

$$z = \frac{\text{point estimate - null value}}{SE} = \frac{0.2 - 0.136}{0.0150} = 4.267$$

(d) **Solutions:** Since we are doing a two-sided test, p-value $= 2 \times P(Z > 4.267)$ and from the table, we could read that the p-value $< 0.0001 < 0.05$. Hence we reject the null hypothesis that the proportion of people that were born in another country is the same for those in the U.S. and China.

3. Suppose you already constructed a 95% confidence interval for the proportion of students that spend their Thanksgiving break at home based on a random sample of 100 students. But now you want to cut the margin of error in half with the same level of confidence. How many students should you survey instead?

Solution: MOE originally is calculated as $z^* \times SE = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$; if we want to cut it in half, with everything else being the same, $\frac{1}{2} \times MOE = \frac{1}{2} \times z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{4n}}$. Hence we need $100 \times 4 = 400$ students if we want to cut MOE in half.

4. In a poll of the percentage of students that approves the selection of a president for student council, the margin of error is calculated to be 4.744%. The approval rate is 48% and this poll is based on a random sample of 300 students. What is the level of confidence that is used? (90%, 95%, 99% or something else?)

Solution: MOE $= z^* \times SE = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z^* \times \sqrt{\frac{0.48 \times 0.52}{300}} = 0.04744$. Solving for z^*, we get that $z^* \approx 1.645$. From the z table, we know that this is constructed based on a 90% confidence level.