MODULE 7 EXAMPLES

TAs: Nihal Reddy
Email: nireddy@ucsd.edu
OH: Thursdays 6-7pm

Slide Credits: Kira Fleischer

PROBLEM #: KEY TOPICS FROM PROBLEM

Problem setup and description.

Question

Key notes from readings/lectures needed to answer the question

Solution: written with as much detail as we expect you to give on your homework sets

PROBLEM 1: CONFIDENCE INTERVALS FOR DIFFERENCE IN PROPORTIONS

Are people who work in non-profit organizations generally happier than those who work in for-profit ones? Data is collected on two random samples. Out of 467 people that work in non-profit organizations, 423 employees said that they are happy. Out of 531 sampled who work in for-profit organizations, 446 said they are happy.

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

$$SE = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Solution: $\hat{p}_1 = \frac{423}{467} = 0.91;\ \hat{p}_2 = \frac{446}{531} = 0.84$

$$SE = \sqrt{\frac{0.91 * 0.09}{467} + \frac{0.84 * 0.16}{531}} = 0.0207$$

PROBLEM 1: CONFIDENCE INTERVALS FOR DIFFERENCE IN PROPORTIONS

Are people who work in non-profit organizations generally happier than those who work in for-profit ones? Data is collected on two random samples. Out of 467 people that work in non-profit organizations, 423 employees said that they are happy. Out of 531 sampled who work in for-profit organizations, 446 said they are happy.

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

1. Independence: The data are independent within and between the two groups.
   - Generally this is satisfied if the data come from two independent random samples or if the data come from a randomized experiment.

2. Success-Failure condition: The success-failure condition holds for both groups, where we check successes and failures in each group separately.
   - $n_1 p_1 \geq 10, n_1(1-p_1) \geq 10; n_2 p_2 \geq 10, n_2(1-p_2) \geq 10$

Solution: Independence within/across groups is met because the samples are both random. The S-F condition is also met:

$$n_1\hat{p}_1 = 423 \geq 10, n_1(1-\hat{p}_1) = 44 \geq 10, n_2\hat{p}_2 = 446 \geq 10, n_2(1-\hat{p}_2) = 85 \geq 10$$

PROBLEM 1: CONFIDENCE INTERVALS FOR DIFFERENCE IN PROPORTIONS

Are people who work in non-profit organizations generally happier than those who work in for-profit ones? Data is collected on two random samples. Out of 467 people that work in non-profit organizations, 423 employees said that they are happy. Out of 531 sampled who work in for-profit organizations, 446 said they are happy.

(c) Construct a 99% two proportion confidence interval for the difference between proportions of happy employees in non-profit and for-profit organizations.

Confidence Interval: $\hat{p}_1 - \hat{p}_2 \pm z^* \times SE$; $z^* = 2.576$ for 99% CI

Solution: From before, we have $\hat{p}_1 = 0.91, \hat{p}_2 = 0.84, SE = 0.0207$. Thus, the CI is:

$$(0.91 - 0.84 \ - [2.576 * 0.0207], 0.91 - 0.84 + [2.576 * 0.0207]) = (1.67\%, 12.33\%)$$

PROBLEM 1: CONFIDENCE INTERVALS FOR DIFFERENCE IN PROPORTIONS

Are people who work in non-profit organizations generally happier than those who work in for-profit ones? Data is collected on two random samples. Out of 467 people that work in non-profit organizations, 423 employees said that they are happy. Out of 531 sampled who work in for-profit organizations, 446 said they are happy.

(d) Interpret the confidence interval computed in context.

Solution: We are 99% confidence that based on these data, the proportion of employees who are happy working in non-profit organizations is between 1.67% and 12.33% higher than the proportion of employees who are happy working in for-profit organizations.

PROBLEM 2: HYPOTHESIS TESTING FOR DIFFERENCE IN PROPORTIONS

Continuing the setting from the previous exercise, we want to test whether the proportion of happy employees at non-profit organizations is the same as those in for-profit organizations.

(a) Define the parameters of interest and write down the proper null and alternative hypotheses.

Solution: Let p_1 denote the proportion of happy employees who work in non-profit organizations and p_2 be the proportion of happy employees who work in for-profit organizations.

$H_0: p_1 - p_2 = 0 \iff$ proportion of happy employees who work in non-profit organizations is the same as those working in for profit organizations

$H_A: p_1 - p_2 \neq 0 \iff$ proportion of happy employees who work in non-profit organizations is different from those working in for profit organizations

PROBLEM 2: HYPOTHESIS TESTING FOR DIFFERENCE IN PROPORTIONS

Continuing the setting from the previous exercise, we want to test whether the proportion of happy employees at non-profit organizations is the same as those in for-profit organizations.

(b) Compute the standard error.

When the null hypothesis is that the proportions are equal, use the pooled proportion to verify the success-failure condition and estimate the standard error:

$$\hat{p}_{pooled} = \frac{\#\ successes}{\#\ cases} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}; \; SE = \sqrt{\frac{\hat{p}_{pooled}(1-\hat{p}_{pooled})}{n_1} + \frac{\hat{p}_{pooled}(1-\hat{p}_{pooled})}{n_2}}$$

Solution: $\hat{p}_{pooled} = \frac{423+446}{467+531} = 0.8707; \; SE = \sqrt{\frac{0.8707*0.1293}{467} + \frac{0.8707*0.1293}{531}} = 0.0213$

PROBLEM 2: HYPOTHESIS TESTING FOR DIFFERENCE IN PROPORTIONS

Continuing the setting from the previous exercise, we want to test whether the proportion of happy employees at non-profit organizations is the same as those in for-profit organizations.

(c) (i) Check the conditions required and (ii) find the z-statistic based on the standard error you calculated.

1. Independence within/across groups: occurs if samples are both random

2. Success-Failure condition: $n_1\hat{p}_{pooled} \geq 10, n_1(1-\hat{p}_{pooled}) \geq 10$;

$n_2\hat{p}_{pooled} \geq 10, n_2(1-\hat{p}_{pooled}) \geq 10$

Solution: Independence within/across groups is met because the samples are both random. The S-F condition is also met: $n_1\hat{p}_{pooled} = 467 * 0.8707 = 406.62 \geq 10$, $n_1(1-\hat{p}_{pooled}) = 60.38 \geq 10$, $n_2\hat{p}_{pooled} = 531 * 0.8707 = 462.34 \geq 10$, $n_2(1-\hat{p}_{pooled}) = 68.66 \geq 10$

PROBLEM 2: HYPOTHESIS TESTING FOR DIFFERENCE IN PROPORTIONS

Continuing the setting from the previous exercise, we want to test whether the proportion of happy employees at non-profit organizations is the same as those in for-profit organizations.

(c) (ii) find the z-statistic based on the standard error you calculated.

$$Z = \frac{point\ estimate - null\ value}{SE}$$

Solution: $Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{SE} = \frac{(0.91 - 0.84) - 0}{0.0213} = 3.286$

PROBLEM 2: HYPOTHESIS TESTING FOR DIFFERENCE IN PROPORTIONS

Continuing the setting from the previous exercise, we want to test whether the proportion of happy employees at non-profit organizations is the same as those in for-profit organizations.

(d) Compute the p-value based on the z-statistic and make a conclusion at the level $\alpha = 0.05$.

Two-sided hypothesis test: p-value $= 2 * P(Z > |z|)$

Solution: p-value $= 2 * P(Z > 3.286) = 2 * (.0005) = .001$

Since the p-value $= 0.001 <$ $\alpha = 0.05$, we reject the null hypothesis that the proportion of happy employees that work in non-profit organizations is the same as those working in for-profit organizations.

PROBLEM 3: MARGIN OF ERROR CALCULATIONS

Suppose you constructed a 99% confidence interval for the proportion of students who spend their Christmas break at home based on a random sample of 50 students. But now you want to reduce the margin of error so that it is one third of the original margin of error with the same level of confidence. **How many students should you survey instead?**

$$MOE = z^* \times SE = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Solution: $\frac{1}{3}MOE = \frac{1}{3} \times z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{9n}}$

Hence we need $9n = 9 * 50 = 450$ students instead.

PROBLEM 4: DETERMINING CONFIDENCE LEVEL

In a survey of the percentage of students that will travel over the summer, the margin of error is calculated to be 10.46%. The proportion of those who will travel over the summer is 56% and this survey is based on a random sample of 150 students. **What is the level of confidence that is used? (90%, 95%, 99% or something else?)**

$$MOE = z^* \times SE = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Solution: $0.1046 = z^* \times \sqrt{\frac{0.56*0.44}{150}}$

Solving for z^*, we get $z^* = 2.58$. From the z-table, we know that this corresponds to a 99% confidence level.