Hypothesis Testing for Two Sample Proportions

Statistics Worksheet · Grade 11–12 / Introductory College Statistics

Name: ________________________________ Date: __________________

Learning Objectives

- Calculate the pooled sample proportion (p-hat c) for two independent samples
- Compute the z-test statistic for a two sample proportion hypothesis test
- Set up null and alternative hypotheses and draw conclusions for two sample proportion tests

Problems

1. In a study, Group 1 has 40 successes out of 200 trials and Group 2 has 60 successes out of 300 trials. Calculate the pooled sample proportion, p-hat c.

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$$

2. For Sample 1: 56 successes out of 2051, and Sample 2: 84 successes out of 2030. Calculate p-hat 1 and p-hat 2 (round to 4 decimal places).

$$\hat{p}_1 = \frac{x_1}{n_1}, \quad \hat{p}_2 = \frac{x_2}{n_2}$$

3. Using the Helsinki Heart Study data (Group 1: 56 heart attacks out of 2051 men on gemfibrozil; Group 2: 84 heart attacks out of 2030 men on placebo), calculate the pooled sample proportion p-hat c. Round to 4 decimal places.

$$\hat{p}_c = \frac{56 + 84}{2051 + 2030}$$

4. A researcher wants to test whether the proportion of smokers who develop lung disease differs between two cities. Write the null hypothesis and the alternative hypothesis for this two-tailed test using p1 for City A and p2 for City B.

$$H_0 : p_1 = p_2 \quad \text{vs.} \quad H_a : p_1 \neq p_2$$

Scan to watch

5. Using p-hat c = 0.0343, n1 = 2051, and n2 = 2030, check all four normality conditions for the two sample proportion test. State whether each condition is satisfied (each product must be greater than 5).

$$n_1\hat{p}_c > 5,\ n_1(1-\hat{p}_c) > 5,\ n_2\hat{p}_c > 5,\ n_2(1-\hat{p}_c) > 5$$

6. A quality control manager samples two production lines. Line 1 has 18 defective items out of 150 and Line 2 has 27 defective items out of 180. Calculate p-hat 1, p-hat 2, and p-hat c.

$$\hat{p}_1 = \frac{18}{150},\quad \hat{p}_2 = \frac{27}{180},\quad \hat{p}_c = \frac{18+27}{150+180}$$

7. Using the quality control data from Problem 6 (p-hat 1 = 0.12, p-hat 2 = 0.15, p-hat c = 0.1364, n1 = 150, n2 = 180), compute the z-test statistic for testing H0: p1 = p2. Round to 2 decimal places.

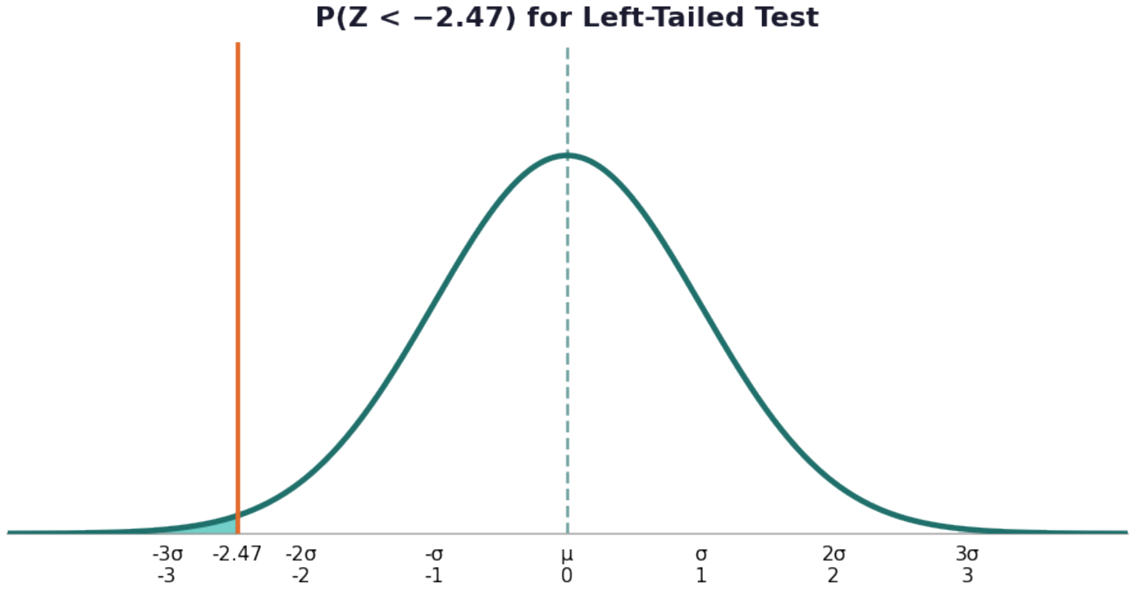
$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_c(1-\hat{p}_c)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

8. Using the Helsinki Heart Study (p-hat 1 = 0.0273, p-hat 2 = 0.0414, p-hat c = 0.0343, n1 = 2051, n2 = 2030), compute the z-test statistic. Round to 2 decimal places.

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_c(1-\hat{p}_c)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

9. For the Helsinki Heart Study, the alternative hypothesis is Ha: p1 < p2 and the significance level is alpha = 0.05. The computed z-test statistic is –2.47. Using the standard normal distribution, find the p-value and state whether you reject or fail to reject the null hypothesis.

Scan to watch



10. In an election poll, 480 out of 1200 voters in County A and 390 out of 900 voters in County B support a new policy. At a significance level of 0.05, test whether the proportion of supporters in County A is less than in County B. State the hypotheses, compute p-hat c and z, find the p-value (left-tailed), and state your conclusion.

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_c(1-\hat{p}_c)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Scan to watch

Hypothesis Testing for Two Sample Proportions — Answer Key

Statistics Worksheet · Grade 11–12 / Introductory College Statistics

Answer Key

1. Answer: p-hat c = 0.20

- Identify values: x1 = 40, x2 = 60, n1 = 200, n2 = 300
- p-hat c = (40 + 60) / (200 + 300) = 100 / 500 = 0.20

2. Answer: p-hat 1 ≈ 0.0273, p-hat 2 ≈ 0.0414

- p-hat 1 = 56 / 2051 ≈ 0.0273
- p-hat 2 = 84 / 2030 ≈ 0.0414

3. Answer: p-hat c ≈ 0.0343

- p-hat c = (56 + 84) / (2051 + 2030) = 140 / 4081
- p-hat c ≈ 0.0343

4. Answer: H0: p1 = p2 (no difference); Ha: p1 ≠ p2 (two-tailed)

- The null hypothesis always states no difference: H0: p1 = p2
- Since we are testing if proportions differ (no direction given), the alternative is two-tailed: Ha: p1 ≠ p2

5. Answer: All four conditions satisfied; normality assumption is met.

- 2051 × 0.0343 ≈ 70.35 > 5 ✓
- 2051 × (1 – 0.0343) ≈ 1980.65 > 5 ✓
- 2030 × 0.0343 ≈ 69.63 > 5 ✓
- 2030 × (1 – 0.0343) ≈ 1960.37 > 5 ✓
- All conditions are satisfied; the normal approximation is valid.

6. Answer: p-hat 1 = 0.12, p-hat 2 = 0.15, p-hat c ≈ 0.1364

- p-hat 1 = 18 / 150 = 0.12
- p-hat 2 = 27 / 180 = 0.15
- p-hat c = (18 + 27) / (150 + 180) = 45 / 330 ≈ 0.1364

7. Answer: z ≈ –0.72

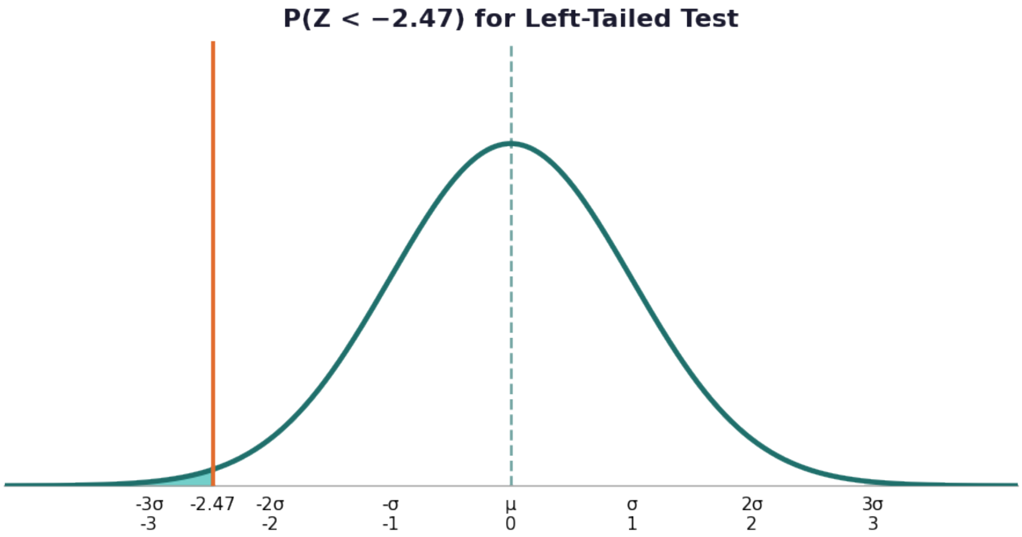
- Numerator: 0.12 – 0.15 = –0.03
- p-hat c × (1 – p-hat c) = 0.1364 × 0.8636 ≈ 0.11779
- 1/n1 + 1/n2 = 1/150 + 1/180 ≈ 0.00667 + 0.00556 = 0.01222
- Denominator: sqrt(0.11779 × 0.01222) ≈ sqrt(0.001439) ≈ 0.03794
- z = –0.03 / 0.03794 ≈ –0.79

8. Answer: z ≈ –2.47

- Numerator: 0.0273 – 0.0414 = –0.0141

- p-hat c × (1 – p-hat c) = 0.0343 × 0.9657 ≈ 0.03312
- 1/n1 + 1/n2 = 1/2051 + 1/2030 ≈ 0.000488 + 0.000493 = 0.000981
- Denominator: sqrt(0.03312 × 0.000981) ≈ sqrt(0.00003249) ≈ 0.005700
- z = –0.0141 / 0.005700 ≈ –2.47

9. Answer: p-value ≈ 0.0068; Reject H0 at α = 0.05



- This is a left-tailed test since Ha: p1 < p2
- z = –2.47
- From the standard normal table: P(Z < –2.47) ≈ 0.0068
- Since p-value (0.0068) < α (0.05), we reject H0
- Conclusion: There is sufficient evidence that the proportion of heart attacks in the gemfibrozil group is significantly less than in the placebo group.

10. Answer: z ≈ –1.47; p-value ≈ 0.0708; Fail to reject H0

- H0: p1 = p2 vs Ha: p1 < p2 (left-tailed test)
- p-hat 1 = 480/1200 = 0.40; p-hat 2 = 390/900 ≈ 0.4333
- p-hat c = (480 + 390) / (1200 + 900) = 870 / 2100 ≈ 0.4143
- p-hat c × (1 – p-hat c) = 0.4143 × 0.5857 ≈ 0.2427
- 1/n1 + 1/n2 = 1/1200 + 1/900 ≈ 0.000833 + 0.001111 = 0.001944
- Denominator: sqrt(0.2427 × 0.001944) ≈ sqrt(0.0004718) ≈ 0.02172
- z = (0.40 – 0.4333) / 0.02172 ≈ –0.0333 / 0.02172 ≈ –1.53
- P(Z < –1.53) ≈ 0.0630; since 0.0630 > 0.05, fail to reject H0
- Conclusion: There is not sufficient evidence at the 0.05 level to conclude that County A has a lower proportion of supporters than County B.

Scan to watch