CEE 260/MIE 273: Probability and Statistics in Civil Engineering

Lecture M5b: Inference for Difference of Two Proportions

Jimi Oke

UMass Amherst

College of Engineering

November 4, 2025

Outline

Today's objectives

- Normality test for difference of two proportions ($p_1 - p_2$)
- Compute CIs for $p_1 - p_2$
- Conduct hypothesis tests for $p_1 - p_2$
- Using the pooled proportion $\hat{p}_{pooled}$

Difference of two proportions

- Earlier, we considered how to perform inference for a *single* population proportion p, using sample estimates $\hat{p}$ (sample proportion) and n (sample size)
- However, there are cases whereby we want to compare proportions from *two groups/populations*
- In such cases, we perform inference on the **difference** of two population proportions p_1 and p_2
- Thus, the parameter of interest is $p_1 - p_2$, and we define the following sample statistics:
  - $\hat{p}_1$: sample proportion for group 1
  - $\hat{p}_2$: sample proportion for group 2
  - $\hat{p}_1 - \hat{p}_2$: difference two sample proportions
  - n_1: sample size of group 1
  - n_2: sample size of group 2

Normality conditions

The difference of two sample proportions $\hat{p}_1 - \hat{p}_2$ can be assumed to follow a normal distribution if:

- The data are obtained from 2 independent random samples (or from a randomized experiment) **[Independence (extended)]**
- The **success-failure condition** holds for both groups separately, i.e.

$$n_1 \hat{p}_1 \geq 10 \tag{1}$$
$$n_1(1 - \hat{p}_1) \geq 10 \tag{2}$$
$$n_2 \hat{p}_2 \geq 10 \tag{3}$$
$$n_2(1 - \hat{p}_2) \geq 10 \tag{4}$$

If these conditions hold, then we can use the normal distribution to find appropriate critical values in order to compute CIs and perform hypothesis tests.

Standard error of $\hat{p}_1 - \hat{p}_2$

In order to compute CIs and perform hypothesis tests for a difference of two proportions $p_1 - p_2$, we need to first find the standard error.
If the normality conditions are satisfied, then the **standard error** of $\hat{p}_1 - \hat{p}_2$ is given by:

$$SE_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \tag{5}$$

where:

- p_1: population proportion for group 1
- p_2: population proportion for group 2
- n_1: sample size of group 1
- n_2: sample size of group 2

In cases where we do not know the population proportions (which is typically the case), we can **approximate** the SE using the sample proportions $\hat{p}_1$ and $\hat{p}_2$. Thus:

$$SE_{\hat{p}_1 - \hat{p}_2} \approx \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \tag{6}$$

CI for difference of two proportions

The confidence interval (CI) for a difference of two proportions is given by:

$$\langle p_1 - p_2 \rangle_{(1-\alpha)} = \underbrace{\hat{p}_1 - \hat{p}_2}_{\text{point estimate}} \pm \underbrace{z^* \times SE}_{\text{margin of error}} \tag{7}$$

where:

$$SE \approx \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \tag{8}$$

$$z^* = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad \text{(critical } Z\text{-score)} \tag{9}$$

Also:

- Φ^{-1} is the inverse CDF function of the standard normal distribution
- α is defined as the **level of significance**
- $1 - \alpha$ is the **level of confidence**

Relationship between confidence and significance levels

For example, if 95% is the desired confidence level, then $1 - \alpha = .95$ and $\alpha = 0.05$

Example 1: CPR Study

Construct and interpret a 95% confidence interval for the difference between two groups of patients in a cardiopulmonary resuscitation (CPR) study. The treatment group received a blood thinner, while the control group did not. Outcome variable of interest: proportion of patients who survived for at least 24 hours.

	Survived	**Died**	**Total**
Treatment (t)	14	26	40
Control (c)	11	39	50
Total	25	65	90

Define:

- p_t: survival rate in treatment group
- p_c: survival rate in control group

Example 1: CPR Study (cont.)

First, we compute the difference of the two sample proportions:

$$\hat{p}_t - \hat{p}_c = \frac{14}{40} - \frac{11}{50} = 0.35 - 0.22 = 0.13$$

Next, we compute the SE:

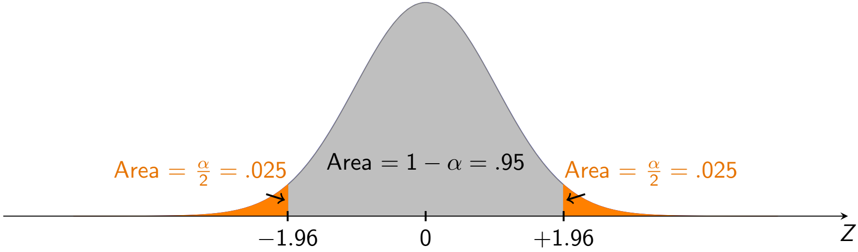
$$\begin{aligned} SE &\approx \sqrt{\frac{\hat{p}_t(1-\hat{p}_t)}{n_t} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_c}} \\ &= \sqrt{\frac{0.35(1-0.35)}{40} + \frac{0.22(1-0.22)}{50}} = 0.095 \end{aligned}$$

Example 1: CPR Study (cont.)

Then we obtain the appropriate critical Z-score z^*.
For a 95% CI, $\alpha = 0.05$. Thus,

$$\begin{aligned} z^* &= \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(1-0.05/2) = \Phi^{-1}(0.975) \\ &= \texttt{norminv}(.975) = 1.96 \end{aligned}$$

Below is the standard normal distribution showing the critical Z-scores corresponding to the desired confidence level of 95%:

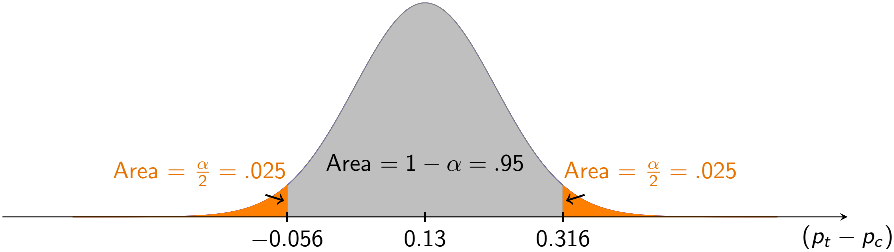


Example 1: CPR Study (cont.)

Thus, the CI is given by:

$$\begin{aligned}\langle p_t - p_c \rangle_{.95} &= 0.13 \pm (1.96 \times 0.095) = 0.13 \pm 0.186 \\ &= (-0.056, 0.316)\end{aligned}$$

Now, we show the corresponding sampling distribution of $p_t - p_c$ with the computed CIs (i.e. the standard normal distribution converted into the actual scale of the point estimate $\hat{p}_1 - \hat{p}_2$)



Example 1: CPR Study (cont.)

Interpretation

Thus, we are 95% confident (*OR* 95% of the time), the difference in survival rate between those who are treated by the blood thinner and those who are not lies between -5.6 and 31.6 percentage points.

Hypothesis testing using critical value

1. State the hypotheses H_0 and H_1
2. Compute the point estimate $\hat{p}_1 - \hat{p}_2$
3. Find the standard error SE
4. Find the test statistic $z = \frac{(\hat{p}_1 - \hat{p}_2) - \Delta_0}{SE}$, where Δ_0 is the null value of the difference
5. Find the critical value(s) z^*
6. Compare the test statistic to the critical value
7. Clearly state the outcome from your hypothesis test
8. Write a final concluding statement in response to the question

Hypothesis testing using p-value

1. State the hypotheses H_0 and H_1
2. Compute the point estimate $\hat{p}_1 - \hat{p}_2$
3. Find the standard error SE
4. Find the p-value
5. Compare the p-value to the level of significance α
6. Clearly state the outcome from your hypothesis test
7. Write a final concluding statement in response to the question

Pooled proportion

If the hypothesis test is to check whether $p_1 = p_2$ or that $p_1 - p_2 = 0$ (null hypothesis), then:

$$H_0 : p_1 - p_2 = \Delta_0 = 0 \quad (10)$$

$$H_1 : p_1 - p_2 \neq 0 \quad (11)$$

In this case, we use the **pooled proportion** $\hat{p}_{pooled}$ to verify the success-failure condition and to estimate the standard error

$$\hat{p}_{pooled} = \frac{\text{number of successes}}{\text{number of successes} + \text{number of failures}} \quad (12)$$

$$= \frac{\text{number of successes}}{\text{total number of cases}} \quad (13)$$

$$= \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2} \quad (14)$$

Thus:

$$SE \approx \sqrt{\frac{\hat{p}_{pooled}(1 - \hat{p}_{pooled})}{n_1} + \frac{\hat{p}_{pooled}(1 - \hat{p}_{pooled})}{n_2}}$$

Example 3: Mammogram study

A study is conducted over a 30-year period with about 90,000 female participants to determine whether mammograms (X-ray procedure to test for breast cancer) are more effective than non-mammogram exams.

	Death from breast cancer?	
	Yes	No
Mammogram (m)	500	44,425
Control (c)	505	44,405

Conduct a hypothesis test determine whether there is no significant evidence to suggest that mammograms are more effective. Use a significance level $\alpha = 0.05$

Hypotheses

Thus, the null hypothesis H_0 is: $p_m - p_c = 0$ (i.e. the null difference $\Delta_0 = 0$).
And the alternative hypothesis H_1 is: $p_m - p_c \neq 0$.

Example 3 (cont.)

In this case the pooled proportion is given by

$$\hat{p}_{pooled} = \frac{\#\text{ patients who died from breast cancer in the entire study}}{\#\text{ patients in the entire study}}$$

$$= \frac{500 + 505}{500 + 44,425 + 505 + 44,405}$$

$$= 0.0112$$

Thus, the SE is given by:

$$SE \approx \sqrt{\frac{\hat{p}_{pooled}(1 - \hat{p}_{pooled})}{n_m} + \frac{\hat{p}_{pooled}(1 - \hat{p}_{pooled})}{n_c}}$$

$$= \sqrt{\frac{0.0112(1 - 0.0112)}{44,925} + \frac{0.0112(1 - 0.0112)}{44,910}} = 0.0007$$

Example 3 (cont.): test statistic

Now we find the point estimate $\hat{p}_m - \hat{p}_c$:

$$\hat{p}_m - \hat{p}_c = \frac{500}{44,925} - \frac{505}{44,910} = -0.00012 \tag{15}$$

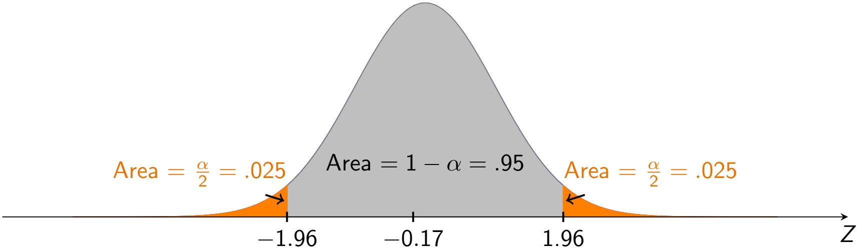
Then, the test statistic is given by:

$$\begin{aligned} z &= \frac{(\hat{p}_m - \hat{p}_c) - \Delta_0}{SE} \\ &= \frac{-0.00012 - 0}{.0007} = -0.17 \end{aligned}$$

Example 3 (cont.): critical value

The critical value z^* in this case is given by:

$$z^* \quad = \quad \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(1-.025) = 1.96$$



Since $-0.17 > -1.96$ and $-0.17 < 1.96$ (i.e. the test statistic is contained within the interval $(-1.96, 1.96)$), then we **fail to reject** the null hypothesis H_0. **Thus, we conclude that there is no evidence to suggest that mammogram exams are significantly more effective than non-mammogram exams**.

Example 3 (cont.)

Next, we will conduct the hypothesis test using p-values (to be continued)