Lecture 10

Quantitative Political Science

Prof. Bisbee

Vanderbilt University

Lecture Date: 2023/10/03

Slides Updated: 2023-10-01

Agenda

1. Finishing up last lecture's example
2. Hypothesis Testing
3. Relation to CIs
4. Two- versus one-tailed tests

Example Time!

- Poll of 1,203 adults between Sep. 15 and 20, 2023 asking about a hypothetical vote choice if the election were held tomorrow, found that 52.5% of respondents indicated they would support Trump, and 47.5% indicated they would support Biden. This marks a reduction in Trump support from a previous tracking poll fielded a week earlier of 1,203 adults who indicated 55.6% support for Trump and 44.4% support for Biden.

- How confident are we that the change in Trump's support over this period is not due to sampling error?

- Parameter we seek is $p_1 - p_2$ where p_1 is Trump's **true** support in the first poll and p_2 is his **true** support in the second poll. Consider the polls as binomial experiments in which Y_1 is the number of "successes" (here, the # favoring Trump) in the first poll and Y_2 is the number of "successes" in the second poll.

- Intuitive estimator: $\hat{p}_1 - \hat{p}_2$. Is this unbiased?

- Calculate estimator's standard errors: $\sqrt{VAR(\hat{p}_1 - \hat{p}_2)} = \sqrt{VAR(\hat{p}_1) + VAR(\hat{p}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Example Time!

- Continuing from the previous example, what is the 95% confidence interval for this estimator?
- Does this interval include zero? How can we interpret that?
- What about the 90% confidence interval? Does it still include zero?
- At what level of confidence would we conclude Trump's support changed between the two surveys?
- **Think**: want to find α (call it α^*) s.t. the *lower bound of the CI is greater than zero*

$$\hat{p}_1 - \hat{p}_2 - z_{\alpha^*/2}\left(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) > 0$$

$$-z_{\alpha^*/2}\left(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) > -(\hat{p}_1 - \hat{p}_2)$$

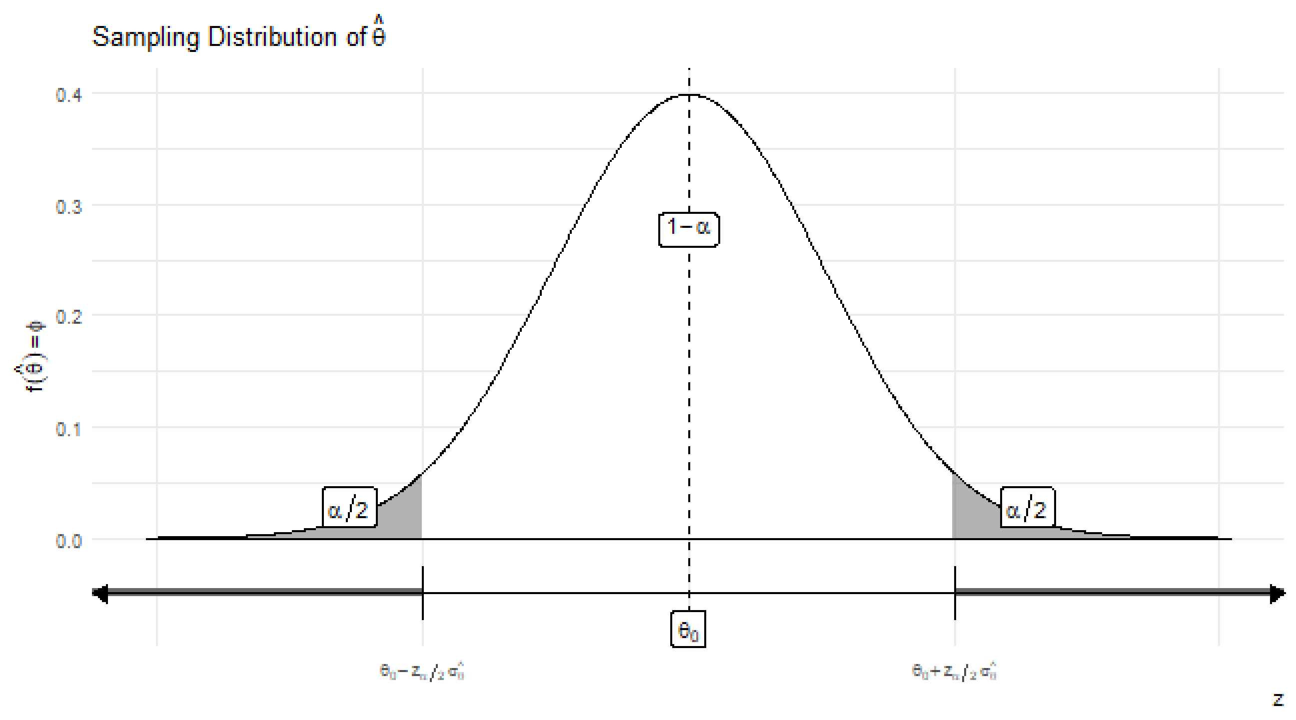
$$z_{\alpha^*/2} < \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Hypothesis Testing

- **Hypothesis Test** consists of four elements:
    1. **Null** hypothesis about a parameter: H_0
    2. **Alternative** hypothesis about the parameter: H_A
    3. **Test statistic** derived from estimator of the parameter
    4. **Rejection region**: range of values of test statistic for which H_0 should be *rejected* in favor of H_A
- Choosing the RR trades off two kinds of errors:
    - **Type I error**: reject H_0 when it is actually true
    - **Type II error**: accept H_0 when H_A is actually true

Type I Error

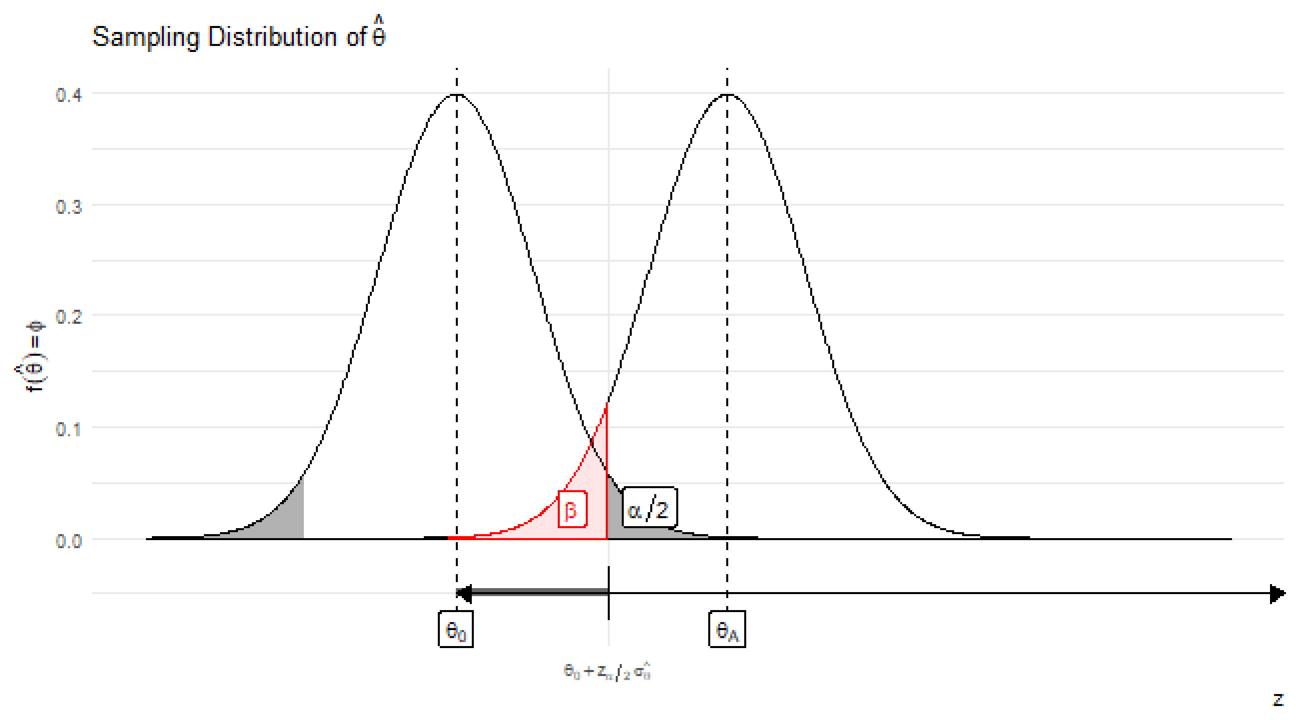
- **Type I error**: reject H_0 when it is actually true
  - What does this look like?



Type I error

- We will (purely by chance):
  - Observe an estimated $\hat{\theta}$ in the *RR* $100 * \alpha\%$ of the time
  - Thus falsely reject the null even though it's true
- This is Type I error!

Type II error



Type II error

- Suppose that the alternative hypothesis is true
- But we always conduct our hypothesis test **under the assumption that the null is true**
- If the sampling distribution of our estimator $\hat{\theta} \sim \mathcal{N}(\theta_A, \sigma_{\hat{\theta}})$, we will mistakenly accept the null $100 * \beta$ % of the time
- Define **power** as $1 - \beta$

$$\begin{aligned} \text{Power} &= 1 - \beta \\ &= 1 - Pr(\text{reject } H_0 | H_A \text{ true}) \\ &= 1 - Pr(\hat{\theta} < \theta_0 + z_{\alpha/2}\sigma_{\hat{\theta}} | \theta = \theta_A) \end{aligned}$$

Type I and II error

- Thus $P(\text{Type I}) = \alpha$ and $P(\text{Type II}) = \beta$
- Ideally, we want the hypothesis test's level of significance to be **low** and its power to be **high**
- Why is this a trade-off?
- Re-evaluate the example:
  - $H_0 : p_1 - p_2 = 0$
  - $H_A : p_1 - p_2 \neq 0$
- Test statistic is $\hat{p}_1 - \hat{p}_2$
- Rejection region is all values of statistic for which we reject H_0 for chosen α
  - I.e., values of $\hat{p}_1 - \hat{p}_2$ where the constructed CI **does not include zero**

Type I and II errror

- Recall our CI: $(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
- We could have rejected H_0 if $(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} > 0$ and concluded Trump's popularity did fall
- Rewriting:

$$(\hat{p}_1 - \hat{p}_2) > z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\frac{(\hat{p}_1 - \hat{p}_2)}{z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} > z_{\alpha/2}$$

Sample Size

- What sample size would have been needed for our difference in sample populations to be statistically significant from zero with 95% confidence? (Assume $n_1 = n_2 = n$)
- Plug in the numbers!
- Power calculation is a crucial tool for determining how big your sample must be to avoid committing Type II error
- (We will come back to this soon)

Relation to CIs

- Walk through a question
  - Conventional wisdom says that $\theta = \theta_0$, but I theorize that $\theta \neq \theta_0$
  - I obtain a point estimate $\hat{\theta} \neq \theta_0$
  - How sure am I that $\theta \neq \theta_0$?
- This is the core language of **hypothesis tests**
- Often $\theta_0 = 0$, but it could be any value
- Regardless, we can fully define the distribution of our estimator if θ_0 is true
- We know that CLT tells us that the standardized version of **any** estimator is $\frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}} \sim \mathcal{N}(0,1)$

Hypothesis Testing

- Remember...4 components!

  1. Null hypothesis H_0
  2. Alternative hypothesis H_A
  3. Test statistic $\hat{\theta}$
  4. Rejection region

- Start by choosing α which is now defined as **the probability of Type I error**
- Then identify the range of values of $\hat{\theta}$ we will observe α percent of the time in repeated sampling, which is our **rejection region**
- If we observe $\hat{\theta}$ in this region, we reject $H_0 : \theta = \theta_0$ in favor of $H_A : \theta \neq \theta_0$
- In practice, we reject H_0 if $\hat{\theta} < \theta_0 - z_{\alpha/2}\sigma_{\hat{\theta}}$ or if $\hat{\theta} > \theta_0 + z_{\alpha/2}\sigma_{\hat{\theta}}$

One-Tailed Hypothesis Test

- What if we have a stronger alternative hypothesis?
  - Our alternative is **signed**
  - Instead of $H_A: \theta \neq \theta_0$, I have theoretical reason to believe $H_A: \theta > \theta_0$
- Again, pick α
- Then look at the standard Normal and identify range of values for $\hat{\theta}$ greater than θ_0 that we will observe α% of the time in repeated sampling
- Beware of cooking the books! Say you have some $\hat{\theta} > \theta_0$, and you make an *ex post* hypothesis that $H_A: \hat{\theta} \geq \theta_0$. This is not based on theory, and looks very suspicious!