

Delta Method for Confidence Interval

Department of Statistics, University of South Carolina

Stat 705: Data Analysis II

Outline

- Review two sample binomial results
- Delta Method

Two sample binomials results

Recall $X \sim \text{Bin}(n_1, p_1)$ and $Y \sim \text{Bin}(n_2, p_2)$. Also this information is often arranged into a 2×2 table:

$n_{11} = X$	$n_{12} = n_1 - X$	n_1
$n_{21} = Y$	$n_{22} = n_2 - Y$	n_2

- $\widehat{RD} = \widehat{p_1} - \widehat{p_2}$
 $\widehat{SE}_{\widehat{RD}} = \sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1} + \frac{\widehat{p_2}(1-\widehat{p_2})}{n_2}}$
- $\widehat{RR} = \frac{\widehat{p_1}}{\widehat{p_2}}$
 $\widehat{SE}_{\log \widehat{RR}} = \sqrt{\frac{(1-\widehat{p_1})}{\widehat{p_1} n_1} + \frac{(1-\widehat{p_2})}{\widehat{p_2} n_2}}$
- $\widehat{OR} = \frac{\widehat{p_1}/(1-\widehat{p_1})}{\widehat{p_2}/(1-\widehat{p_2})} = \frac{n_{11} n_{22}}{n_{12} n_{21}}$
 $\widehat{SE}_{\log \widehat{OR}} = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$
- $\text{CI} = \text{Estimate} \pm Z_{1-\alpha/2} \text{SE}_{\text{Est}}$

Motivation for the Delta method

- If $\hat{\theta}$ is close to θ then

$$\frac{f(\hat{\theta}) - f(\theta)}{\hat{\theta} - \theta} \approx f'(\hat{\theta})$$

- So,

$$\frac{f(\hat{\theta}) - f(\theta)}{f'(\hat{\theta})} \approx \hat{\theta} - \theta$$

- Therefore,

$$\frac{f(\hat{\theta}) - f(\theta)}{f'(\hat{\theta})\hat{SE}_{\hat{\theta}}} \approx \frac{\hat{\theta} - \theta}{\hat{SE}_{\hat{\theta}}}$$

Standard errors

- Delta method can be used to obtain large sample standard errors
- Formally, the delta method states that if

$$\frac{\hat{\theta} - \theta}{\hat{SE}_{\hat{\theta}}} \rightarrow N(0, 1)$$

- Then

$$\frac{f(\hat{\theta}) - f(\theta)}{f'(\hat{\theta})\hat{SE}_{\hat{\theta}}} \rightarrow N(0, 1)$$

- Asymptotic mean of $f(\hat{\theta})$ is $f(\theta)$
- Asymptotic standard error of $f(\hat{\theta})$ can be estimated with $f'(\hat{\theta})\hat{SE}_{\hat{\theta}}$

Example

- $\theta = p_1$
- $\hat{\theta} = \hat{p}_1$
- $\hat{SE}_{\hat{\theta}} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}$
- $f(x) = \log(x)$
- $f'(x) = \frac{1}{x}$
- $\frac{\hat{\theta} - \theta}{\hat{SE}_{\hat{\theta}}} \rightarrow N(0, 1)$ by CLT
- Then $\hat{SE}_{\log \hat{p}_1} = f'(\hat{\theta}) \hat{SE}_{\hat{\theta}}$

$$\begin{aligned} &= \frac{1}{\hat{p}_1} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}} = \sqrt{\frac{(1-\hat{p}_1)}{\hat{p}_1 n_1}} \\ &\frac{\log \hat{p}_1 - \log p_1}{\sqrt{\frac{(1-\hat{p}_1)}{\hat{p}_1 n_1}}} \rightarrow N(0, 1) \end{aligned}$$

Putting it all together

- Asymptotic standard error

$$\begin{aligned}\text{Var}(\log \widehat{RR}) &= \text{Var}\{\log(\widehat{p}_1/\widehat{p}_2)\} \\ &= \text{Var}(\log \widehat{p}_1) + \text{Var}(\log \widehat{p}_2) \\ &\approx \frac{(1 - \widehat{p}_1)}{\widehat{p}_1 n_1} + \frac{(1 - \widehat{p}_2)}{\widehat{p}_2 n_2}\end{aligned}$$

- The last line following from the delta method
- The approximation requires large sample sizes
- The delta method can be used similarly for the log odds ratio