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 $X \sim Bin(n_1, p_1)$ and $Y \sim Bin(n_2, p_2)$. 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

- $\widehat{RD} = \widehat{p_1 - p_2}$
  $\hat{SE}_{\hat{RD}} = \sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}$
- $\widehat{RR} = \frac{\widehat{p}_1}{\widehat{p}_2}$
  $\hat{SE}_{\log \hat{RR}} = \sqrt{\frac{(1-\widehat{p}_1)}{\widehat{p}_1 n_1} + \frac{(1-\widehat{p}_2)}{\widehat{p}_2 n_2}}$
- $\widehat{OR} = \frac{\widehat{p}_1/(1-\widehat{p}_1)}{\widehat{p}_2/(1-\widehat{p}_2)} = \frac{n_{11} n_{22}}{n_{12} n_{21}}$
  $\hat{SE}_{\log \hat{OR}} = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$
- $\text{CI} = \text{Estimate} \pm Z_{1-\alpha/2} \text{SE}_{\text{Est}}$

Motivation for the Delta method

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

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

- So,

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

- Therefore,

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

Standard errors

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

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

- Then

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

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

Example

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

$$= \frac{1}{\widehat{p_1}}\sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1}} = \sqrt{\frac{(1-\widehat{p_1})}{\widehat{p_1}n_1}}$$

$$\frac{\log \widehat{p_1} - \log p_1}{\sqrt{\frac{(1-\widehat{p_1})}{\widehat{p_1}n_1}}} \to N(0,1)$$

Putting it all together

- Asymptotic standard error

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

- 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