

Lecture 19

Ciprian M Crainiceanu

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
Johns Hopkins University

September 8, 2020

Table of contents

① Table of contents

② Outline

③ Relative measures

④ The relative risk

⑤ The odds ratio

- 1 Define relative risk
- 2 Odds ratio
- 3 Confidence intervals

Motivation

- Consider a randomized trial where 40 subjects were randomized (20 each) to two drugs with the same active ingredient but different expedients
- Consider counting the number of subjects with side effects for each drug

	Side		total
	Effects	None	
Drug A	11	9	20
Drug B	5	15	20
Total	16	14	40

Comparing two binomials

- Let $X \sim \text{Binomial}(n_1, p_1)$ and $\hat{p}_1 = X/n_1$
- Let $Y \sim \text{Binomial}(n_2, p_2)$ and $\hat{p}_2 = Y/n_2$
- We also use the following notation:

$n_{11} = X$	$n_{12} = n_1 - X$	$n_1 = n_{1+}$
$n_{21} = Y$	$n_{22} = n_2 - Y$	$n_2 = n_{2+}$
n_{2+}	n_{+2}	

- Last time, we considered the absolute change in the proportions, what about relative changes?
- Relative changes are often of more interest than absolute, eg when both proportions are small
- The **relative risk** is defined as p_1/p_2
- The natural estimator for the relative risk is

$$\hat{RR} = \frac{\hat{p}_1}{\hat{p}_2} = \frac{X/n_1}{Y/n_2}$$

- The standard error for $\log \hat{RR}$ is

$$\hat{SE}_{\log \hat{RR}} = \left(\frac{(1 - p_1)}{p_1 n_1} + \frac{(1 - p_2)}{p_2 n_2} \right)^{1/2}$$

- Exponentiate the resulting interval to get an interval for the RR

- The **odds ratio** is defined as

$$\frac{\text{Odds of SE Drug A}}{\text{Odds of SE Drug B}} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{p_1(1-p_2)}{p_2(1-p_1)}$$

- The sample odds ratio simply plugs in the estimates for p_1 and p_2 , this works out to have a convenient form

$$\hat{OR} = \frac{\hat{p}_1/(1-\hat{p}_1)}{\hat{p}_2/(1-\hat{p}_2)} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

(cross product ratio)

- The standard error for $\log \hat{OR}$ is

$$\hat{SE}_{\log \hat{OR}} = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

- Exponentiate the resulting interval to obtain an interval for the OR

Some comments

- Notice that the sample and true odds ratios do not change if we transpose the rows and the columns
- For both the OR and the RR, taking the logs helps with adherence to the error rate
- Of course the interval for the log RR or log OR is obtained by taking

$$\text{Estimate} \pm Z_{1-\alpha/2} SE_{\text{Estimate}}$$

- Exponentiating yields an interval for the OR or RR
- Though logging helps, these intervals still don't perform altogether that well

Example - RR

- For the relative risk, $\hat{p}_A = 11/20 = .55$, $\hat{p}_B = 5/20 = .25$
- $\hat{RR}_{A/B} = .55/.25 = 2.2$
- $\hat{SE}_{\log \hat{RR}_{A/B}} = \sqrt{\frac{1-.55}{.55 \times 20} + \frac{1-.25}{.25 \times 20}} = .44$
- Interval for the log RR:
 $\log(2.2) \pm 1.96 \times .44 = [-.07, 1.65]$
- Interval for the RR: $[.93, 5.21]$

Example - OR

- $\hat{OR}_{A/B} = \frac{11 \times 15}{9 \times 5} = 3.67$
- $\hat{SE}_{\log \hat{OR}_{A/B}} = \sqrt{\frac{1}{11} + \frac{1}{9} + \frac{1}{5} + \frac{1}{15}} = .68$
- Interval for log OR: $\log(3.67) \pm 1.96 \times .68 = [-.04, 2.64]$
- Interval for the OR: $[.96, 14.01]$

Example - RD

- For the risk difference
$$\hat{RD}_{A-B} = \hat{p}_A - \hat{p}_B = .55 - .25 = .30$$
- $\hat{SE}_{\hat{RD}_{A-B}} = \sqrt{\frac{.55 \times .45}{20} + \frac{.25 \times .75}{20}} = .15$
- Interval: $.30 \pm 1.96 \times .15 = [.15, .45]$



