

BIOS 6244 Analysis of Categorical Data
Assignment 2 Solutions

- (1) Work Exercise 2.1, p. 45. For part (b), find an approximate 95% CI for the risk difference, RR, and OR.

Solution

	MI Death		
	Yes	No	
Aspirin	18	658	676
No Aspirin	28	656	684

a. Risk difference = $\frac{18}{676} - \frac{28}{684} = .0266 - .0409 = -.0143$

Interpretation: There was a 1.4% decrease in the risk of fatal MI among those who took aspirin.

Relative risk = $\frac{.0266}{.0409} = .650$

Interpretation: The risk of fatal MI was reduced by 35% among those who took aspirin.

Odds ratio = $\frac{(18)(656)}{(28)(658)} = .641$

Interpretation: The odds of fatal MI were reduced by approximately 36% among those who took aspirin.

- b. Approximate 95% CI($\pi_1 - \pi_2$)

$$\begin{aligned}
 (p_1 - p_2) \pm 1.96 \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} &= \left(\frac{18}{676} - \frac{28}{684} \right) \pm 1.96 \sqrt{\frac{(.0266)(1-.0266)}{676} + \frac{(.0409)(1-.0409)}{684}} \\
 &= -.0143 \pm 1.96(.0098) = -.0143 \pm .0192 \\
 &= (-.034, .005)
 \end{aligned}$$

Exercise (1), cont.

Since $0 \in \text{CI}(\pi_1 - \pi_2)$, we conclude that there is no significant difference, i.e., that taking aspirin is not associated with a change in risk of fatal MI.

Approximate 95% CI[log(RR)] (from Exercise 2.12)

$$\begin{aligned} \log(\widehat{\text{RR}}) \pm 1.96 \sqrt{\frac{1-p_1}{N_1 p_1} + \frac{1-p_2}{N_2 p_2}} &= \log\left(\frac{.0266}{.0409}\right) \pm 1.96 \sqrt{\frac{1-.0266}{(676)(.0266)} + \frac{1-.0409}{(684)(.0409)}} \\ &= -.4302 \pm 1.96(.2973) \\ &= -.4302 \pm .5827 = (-1.0129, .1525). \end{aligned}$$

Back-transforming, we obtain an approximate 95% CI(RR): $(e^{-1.0129}, e^{.1525}) = (.363, 1.165)$.

Since $1 \in \text{CI}(\text{RR})$, we conclude that taking aspirin is not associated with a change in risk of fatal MI.

Approximate 95% CI[log(OR)]

$$\begin{aligned} \log(\widehat{\text{OR}}) \pm 1.96 \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}} &= \log(.6409) \pm 1.96 \sqrt{\frac{1}{18} + \frac{1}{658} + \frac{1}{28} + \frac{1}{656}} \\ &= -.4449 \pm 1.96(.3071) = -.4449 \pm .6019 \\ &= (-1.0468, .1570) \end{aligned}$$

Back-transforming, we obtain an approximate 95% CI(OR): $(e^{-1.0468}, e^{.1570}) = (.351, 1.170)$.

Since $1 \in \text{CI}(\text{OR})$, we conclude that taking aspirin is not associated with a change in risk of fatal MI.

Overall conclusion:

Taking all of these results into account, we conclude that there appears to be no protective effect of aspirin with regard to fatal MI among people who have already suffered a stroke.

(2) Work Exercise 2.2, pp. 45-46 in our textbook.

Solution

a. Risk difference = $.001304 - .000121 = .001$

Interpretation: smoking is associated with an increased risk of lung cancer of .1% among women over the age of 35

$$\text{Relative risk} = \frac{.001304}{.000121} = 10.78$$

Interpretation: smoking increases risk of lung cancer by a factor of almost 11 among women over the age of 35.

The RR is more informative since the risk difference makes it appear that there is no association between smoking and lung cancer.

b. Odds ratio = $\frac{(.001304)(1 - .000121)}{(.000121)(1 - .001304)} = 10.79$

Interpretation: smoking increases the odds of lung cancer by a factor of almost 11 among women over the age of 35.

The RR and OR are almost identical because the proportion of “diseased” is so close to 0 in both the exposed and unexposed groups.

(3) Work Exercise 2.10, p. 47 in our textbook.

Solution

a. Risk difference = $.00140 - .00010 = .00130$ lung cancer
 $= .00669 - .00413 = .00256$ heart disease

Relative risk = $\frac{.00140}{.00010} = 14.00$ lung cancer
 $= \frac{.00669}{.00413} = 1.62$ heart disease

$$\begin{aligned} \text{Odds ratio} &= \frac{(.0014)(1-.0001)}{(.0001)(1-.0014)} = 14.02 && \text{lung cancer} \\ &= \frac{(.00669)(1-.00413)}{(.00413)(1-.00669)} = 1.62 && \text{heart disease} \end{aligned}$$

The risk difference does not indicate much of an increase in risk for either lung cancer or heart disease.

The RR & OR both indicate a huge increase in risk for lung cancer but not for heart disease.

- b. If smoking were eliminated, there would be 256 fewer deaths per year due to heart disease per 100,000 population. There would only be 130 fewer deaths due to lung cancer. Therefore, heart disease is more strongly related to cigarette smoking in terms of the yearly reduction in deaths per 100,000 population.

NOTE: The RR and OR are generally preferred vs. the risk difference when examining the association between “exposure” and “disease.” However, there are circumstances (as in Exercise 2.10) when the risk difference is preferred to the RR and OR.