STA 6384, Report 3.6

Carson Slater *Baylor University*

Problem: Work problem 3.4, p 104 of Agresti

For Table 2.10 on seat-belt use and results of auto accidents, find and interpret 95% confidence intervals for the conceptual population (a) odds ratio, (b) difference of proportions, and (c) relative risk.

The data on seat-belt use and injury severity is summarized as follows:

Seat-Belt Use	Fatal Injury	Nonfatal Injury	Total
No	1085	55,623	56,708
Yes	703	441,239	441,942

Let Group 1 refer to those not wearing a seat belt and Group 2 refer to those wearing a seat belt. The outcome of interest is a fatal injury.

- $n_{11} = 1085$ (Fatal, No Belt), $n_{12} = 55,623$ (Nonfatal, No Belt), $n_1 = 56,708$
- $n_{21} = 703$ (Fatal, Yes Belt), $n_{22} = 441,239$ (Nonfatal, Yes Belt), $n_2 = 441,942$

Let π_1 and π_2 be the population proportions of fatal injuries for Group 1 and Group 2, respectively. The sample proportions of fatal injuries are:

$$\hat{\pi}_1 = \frac{n_{11}}{n_1} = \frac{1085}{56708} \approx 0.019133$$

$$\hat{\pi}_2 = \frac{n_{21}}{n_2} = \frac{703}{441942} \approx 0.001591$$

(a) 95% Confidence Interval for the Odds Ratio (OR)

The sample odds ratio ($\hat{\theta}$) compares the odds of a fatal injury in the 'No Seat-Belt' group to the 'Yes Seat-Belt' group.

$$\hat{\theta} = \frac{\text{odds}_1}{\text{odds}_2} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{1085/55623}{703/441239} = \frac{1085 \times 441239}{55623 \times 703} \approx 12.243$$

The 95% confidence interval for the odds ratio is constructed on the log scale. The standard error of the log odds ratio is:

$$\text{SE}(\log \hat{\theta}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}} = \sqrt{\frac{1}{1085} + \frac{1}{55623} + \frac{1}{703} + \frac{1}{441239}} \approx 0.04863$$

The 95% CI for $\log(\theta)$ is $\log(\hat{\theta}) \pm 1.96 \cdot \text{SE}(\log \hat{\theta})$:

$$\log(12.243) \pm 1.96 \times 0.04863 \implies 2.505 \pm 0.0953 \implies (2.4097, 2.6003)$$

Exponentiating the endpoints gives the 95% CI for the odds ratio:

$$(\exp(2.4097), \exp(2.6003)) \implies (11.13, 13.47)$$

Interpretation: We are 95% confident that the odds of a fatal injury for an individual not wearing a seat belt are between **11.13 and 13.47 times** the odds for an individual wearing one. Since the interval does not include 1, there is a statistically significant association between seat-belt use and the odds of a fatal injury.

(b) 95% Confidence Interval for the Difference of Proportions

The sample difference in the proportions of fatal injuries is:

$$\hat{\pi}_1 - \hat{\pi}_2 = 0.019133 - 0.001591 = 0.017542$$

The standard error for the difference of proportions is:

$$\text{SE}(\hat{\pi}_1 - \hat{\pi}_2) = \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}} = \sqrt{\frac{0.019133(0.980867)}{56708} + \frac{0.001591(0.998409)}{441942}} \approx 0.000578$$

The 95% CI for the difference $\pi_1 - \pi_2$ is $(\hat{\pi}_1 - \hat{\pi}_2) \pm 1.96 \cdot \text{SE}(\hat{\pi}_1 - \hat{\pi}_2)$:

$$0.017542 \pm 1.96 \times 0.000578 \implies 0.017542 \pm 0.001133 \implies (0.0164, 0.0187)$$

Interpretation: We are 95% confident that the proportion of fatal injuries is between **0.0164 and 0.0187** (or 1.64% and 1.87%) higher for people not wearing seat belts compared to those who do. As the interval does not contain 0, the difference is statistically significant.

(c) 95% Confidence Interval for the Relative Risk (RR)

The sample relative risk (or risk ratio) compares the probability of a fatal injury between the two groups.

$$\hat{RR} = \frac{\hat{\pi}_1}{\hat{\pi}_2} = \frac{0.019133}{0.001591} \approx 12.026$$

Similar to the odds ratio, the confidence interval for the relative risk is found via the log scale. The standard error of the log relative risk is:

$$\text{SE}(\log \hat{RR}) = \sqrt{\frac{1-\hat{\pi}_1}{n_{11}} + \frac{1-\hat{\pi}_2}{n_{21}}} = \sqrt{\frac{1-0.019133}{1085} + \frac{1-0.001591}{703}} \approx 0.04821$$

The 95% CI for $\log(RR)$ is $\log(\hat{RR}) \pm 1.96 \cdot \text{SE}(\log \hat{RR})$:

$$\log(12.026) \pm 1.96 \times 0.04821 \implies 2.487 \pm 0.0945 \implies (2.3925, 2.5815)$$

Exponentiating the endpoints gives the 95% CI for the relative risk:

$$(\exp(2.3925), \exp(2.5815)) \implies (10.94, 13.22)$$

Interpretation: We are 95% confident that the risk of a fatal injury in an auto accident for someone not wearing a seat belt is between **10.94 and 13.22 times** the risk for someone wearing a seat belt. Since this interval does not include 1, the increased risk is statistically significant.