Discrete Response Model
Lecture 1

datascience@berkeley

Odd Ratios

MLE

$$OR = \frac{odds_1}{odds_2} = \frac{\hat{\pi}_1(1-\hat{\pi}_2)}{\hat{\pi}_2(1-\hat{\pi}_1)}$$

$$= \frac{w_1/n_1(1-w_2/n_2)}{w_2/n_2(1-w_1/n_1)} = \frac{w_1(n_2-w_2)}{w_2(n_1-w_1)}$$

The estimate is a product of the counts on the "diagonal" (top left to bottom right) of the contingency table divided by a product of the counts on the off diagonal.

Interpretation

$$OR = \frac{odds_1}{odds_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)} = \frac{\pi_1 (1 - \pi_2)}{\pi_2 (1 - \pi_1)}$$

- Remember that Odd ratios is the ratio of two odds, comparing the odds of success relative to the odds of failure.
- The estimated odds of a success are OR times as large as in Group 1 than in Group 2.
- The estimated odds of a success are 1/OR times as large as in Group 2 than in Group 1.

Interpretation

$$\frac{(1-\hat{\pi}_1)/\hat{\pi}_1}{(1-\hat{\pi}_2)/\hat{\pi}_2} = \frac{\hat{\pi}_2(1-\hat{\pi}_1)}{\hat{\pi}_1(1-\hat{\pi}_2)}$$

- Consider the odds of a failure $(1-\pi_1)/\ \pi_1$, so the ratio of Group 1 to Group 2 becomes.

- The estimated odds of a failure are 1/OR times as large as in Group 1 than in Group 2 and OR times as large as in Group 2 than in Group 1.

MLE

- Because OR is a maximum likelihood estimate, we can use the "usual" properties of them to find the confidence interval.
- However, using the log(OR) often works better (i.e., its distribution is closer to being a normal distribution).

$$\mathrm{Var}(\log(\mathrm{OR})) = \frac{1}{w_1} + \frac{1}{n_1 - w_1} + \frac{1}{w_2} + \frac{1}{n_2 - w_2}.$$

The $(1 - \alpha)100\%$ Wald confidence interval for log(OR) is

$$\log(\mathrm{OR}) \pm Z_{1-\alpha/2}\sqrt{\frac{1}{w_1} + \frac{1}{n_1 - w_1} + \frac{1}{w_2} + \frac{1}{n_2 - w_2}}$$

The $(1 - \alpha)100\%$ Wald confidence interval for OR is

$$\exp\left[\log(\mathrm{OR}) \pm Z_{1-\alpha/2}\sqrt{\frac{1}{w_1} + \frac{1}{n_1 - w_1} + \frac{1}{w_2} + \frac{1}{n_2 - w_2}}\right]$$

Example: Larry Bird's Free Throw

		Second		
		Made	Missed	Total
First	Made	251	34	285
	Missed	48	5	53
	Total	299	39	338

$$OR = \frac{w_1(n_2 - w_2)}{w_2(n_1 - w_1)} = \frac{251 * 5}{48 * 34} = 0.7690.$$

Interpretation:

- The estimated odds of a made second free throw are **0.7690 times as large** when the first free throw is made than when the first free throw is missed.
- The estimated odds of a made second free throw are **1/0.7690 = 1.3 times as large** when the first free throw is missed than when the first free throw is made.

Example: Larry Bird's Free Throw

- In practice, present only one of these interpretations.
- Prefer the second interpretation and could rephrase it as "The estimated odds of a made second free throw are **30% larger** when the first free throw is missed than when the first free throw is made."

Incorrect Interpretation:

- "The estimated odds of a made second free throw are 1.3 times as **likely** … " is incorrect because "likely" means probabilities are being compared.
- Replacing "odds" with "probability" in any correct interpretation; remember, "odds" are not the same as probabilities.
- "The estimated odds are 1.3 times higher …" is incorrect because 1.3 means 30% times higher, not 130%. [Review the relative risk interpretation discussion if needed.]

Berkeley SCHOOL OF INFORMATION