Module 2 – Section 3

Relative Risk and the Odds Ratio

Overview

- Relative Risk
- Odds Ratio

Variables

- Variable 2 = Response Variable
  - $J = 2$ categories
    - Success/Failure
    - Category of Interest/Not Category of Interest
- Variable 1 = Grouping Variable
  - $I = 2$ groups (categories)

Population Proportions

- p_1 = probability of success in group 1
- p_2 = probability of success in group 2

Comparing p_1 and p_2

- Difference ignores relative size of two proportions:

p_1	p_2	$p_1 - p_2$	${p_1}/{p_2}$
0.01	0.001	0.009	10
0.05	0.041	0.009	1.220
0.10	0.091	0.009	1.099
0.25	0.241	0.009	1.037
0.41	0.401	0.009	1.022

Comparing p_1 and p_2

- Same difference in values, but ratio is different
  - 0.01 vs. 0.001 – probability of success in group 1 is 10 times the probability of success in group 2.
  - 0.41 vs. 0.401 – probability of success in group 1 is 1.022 times the probability of success in group 2.

Population Relative Risk

- Ratio of probability of success between two groups

$$RR = \frac{p_1}{p_2}$$

- If $p_1 = p_2$, $RR = 1$
- If $p_1 > p_2$, $RR > 1$
- If $p_1 < p_2$, $RR < 1$

Population Relative Risk

- The convention is to calculate and interpret relative risk so that it is greater than 1.
- If $p_1 < p_2$, compare probability of group 2 to group 1

$$RR = \frac{p_2}{p_1}$$

Estimating Population Relative Risk

- Ratio of observed probability of success between two groups

$$\widehat{RR} = \frac{\hat{p}_1}{\hat{p}_2} = \frac{\frac{Y_{11}}{n_1}}{\frac{Y_{21}}{n_2}}$$

	Response Variable		
Explanatory Variable	Success	Failure	Total
Group 1	Y_{11}	Y_{12}	n_1
Group 2	Y_{21}	Y_{22}	n_2
Total	$Y_{.1}$	$Y_{.2}$	n

Ex. Prostate Cancer

- $\hat{p}_1 = 0.0461, \hat{p}_2 = 0.0891$

$$\widehat{RR} = \frac{\hat{p}_1}{\hat{p}_2} = \frac{0.0461}{0.0891} = 0.5174$$

- Interpret value as $\frac{\hat{p}_2}{\hat{p}_1} = \frac{0.0891}{0.0461} = 1.933$

Ex. Interpretation

- The observed risk (or probability) of a person with prostate cancer dying from prostate cancer if the person does not have surgery is 1.933 times the observed risk (or probability) of a person with prostate cancer dying from prostate cancer if the person has surgery.

Estimating Population Relative Risk

- $\widehat{RR}$ is a statistic – will vary between samples.
- Estimate by calculating confidence interval
- Two Steps:
  1. Calculate confidence interval for $\ln(RR)$
  2. Apply exponential function to endpoints to obtain confidence interval for RR

Confidence Interval for $\ln(RR)$

- Using $\hat{p}_1$ and $\hat{p}_2$

$$\ln(\widehat{RR}) \pm z_{1-\alpha/2}\sqrt{\frac{1-\hat{p}_1}{n_1\hat{p}_1}+\frac{1-\hat{p}_2}{n_2\hat{p}_2}}$$

- Using data from contingency table

$$\ln(\widehat{RR}) \pm z_{1-\alpha/2}\sqrt{\frac{Y_{12}}{n_1Y_{11}}+\frac{Y_{22}}{n_2Y_{21}}}$$

Confidence Interval for RR

- Take exponential function of endpoints to return values to original scale.
- If CI for $\ln(RR) = (a, b)$, then CI for RR is:

$$(\exp(a), \exp(b))$$

Ex. Prostate Cancer

- 95% CI for $\ln(RR)$

$$\ln(0.5174) \pm 1.96\sqrt{\frac{0.9539}{16} + \frac{0.9109}{31}} = (-1.2437, -0.0742)$$

- 95% CI for RR

$$(\exp(-1.2437), \exp(-0.0742)) = (0.2883, 0.9285)$$

Ex. Prostate Cancer

- Since values are less than 1, we will interpret reciprocal CI.

$$\left(\frac{1}{0.9285}, \frac{1}{0.2883}\right) = (1.077, 3.469)$$

Ex. Interpretation

- In this population, we are 95% confident the probability (risk) of a person with prostate cancer dying from prostate cancer if the person does not have surgery is between 1.077 and 3.469 times the probability (risk) of a person with prostate cancer dying from prostate cancer if the person has surgery.

Ex. Note

- It is not likely that a person with prostate cancer will die from prostate cancer in either group.
- A high **relative** risk does not indicate a high risk for that group.

Population Proportion and Odds

- p = probability of success
- Population odds of success

$$\frac{p}{1-p}$$

Population Odds

- $p = 0.5$
  - Odds = 1
- $p > 0.5$
  - Odds > 1
- $p < 0.5$
  - Odds < 1

Population Odds

- Given Odds, you can find p:

$$p = \frac{\text{Odds}}{\text{Odds} + 1}$$

- Ex. Odds = 2 means $p = \frac{2}{3}$

Population Odds

p	$1 - p$	Odds
0.1	0.9	1/9
0.25	0.75	1/3
0.5	0.5	1
0.75	0.25	3
0.9	0.1	9

Population Odds Ratio

- Ratio of population odds of success for group 1 to group 2.

$$\phi = \frac{\left(\frac{p_1}{1 - p_1}\right)}{\left(\frac{p_2}{1 - p_2}\right)}$$

Population Odds Ratio

- $\phi = 1$
  - The odds of success is the same for the two groups.
  - Implies $p_1 = p_2$
- $\phi > 1$
  - The odds of success is larger for group 1 than for group 2.
  - Implies $p_1 > p_2$

Population Odds Ratio

- $\phi < 1$
  - The odds of success is smaller for group 1 than for group 2.
  - Implies $p_1 < p_2$
- Similar to relative risk, we generally interpret odds ratios < 1 by taking the reciprocal.

Population Odds Ratio

- Notice value of ϕ does not indicate the actual values of p_1 and p_2.
- Implies **relative** value, not actual value.

Population Odds Ratio

p_1	p_2	Odds 1	Odds 2	Odds Ratio
1/11	1/21	0.1	0.05	2
1/6	1/11	0.2	0.1	2
0.2	1/9	0.25	0.125	2
1/3	0.2	0.5	0.25	2
0.5	1/3	1	0.5	2

p_1	p_2	Odds 1	Odds 2	Odds Ratio
2/3	0.5	2	1	2
0.8	2/3	4	2	2
6/7	0.75	6	3	2
8/9	0.8	8	4	2
10/11	5/6	10	5	2

Estimating Population Odds Ratio

- Ratio of observed odds of success between two groups.

$$\hat{\phi} = \frac{\left(\frac{\hat{p}_1}{1-\hat{p}_1}\right)}{\left(\frac{\hat{p}_2}{1-\hat{p}_2}\right)} = \frac{\frac{Y_{11}}{Y_{12}}}{\frac{Y_{21}}{Y_{22}}} = \frac{Y_{11}Y_{22}}{Y_{12}Y_{21}}$$

	Response Variable		
Explanatory Variable	Success	Failure	Total
Group 1	Y_{11}	Y_{12}	n_1
Group 2	Y_{21}	Y_{22}	n_2
Total	$Y_{.1}$	$Y_{.2}$	n

Ex. Survey to Doctors

- $\hat{p}_1 = 0.5122,\ \hat{p}_2 = 0.5259$

$$\hat{\phi} = \frac{\left(\frac{0.5122}{0.4878}\right)}{\left(\frac{0.5259}{0.4741}\right)} = 0.9466$$

Ex. Interpretation

- Since $\hat{\phi} < 1$, interpret reciprocal $\left(\frac{1}{0.9466}\right) = 1.0564$

- In our data, the odds of returning the survey for the group that did not receive the letter is 1.0564 times the odds of returning the survey for the group that received the letter.

Estimating Population Odds Ratio

- Sample odds ratio $\hat{\phi}$ varies between samples.
- Estimate by calculating confidence interval
- Two Steps:
  1. Calculate confidence interval for $\ln(\phi)$
  2. Apply exponential function to endpoints to obtain confidence interval for ϕ

Confidence Interval for $\ln(\phi)$

- Using $\hat{p}_1$ and $\hat{p}_2$

$$SE\left(\ln\left(\hat{\phi}\right)\right) = \sqrt{\frac{1}{n_1\hat{p}_1} + \frac{1}{n_1(1-\hat{p}_1)} + \frac{1}{n_2\hat{p}_2} + \frac{1}{n_2(1-\hat{p}_2)}}$$

- Using data from Contingency Table

$$SE\left(\ln\left(\hat{\phi}\right)\right) = \sqrt{\frac{1}{Y_{11}} + \frac{1}{Y_{12}} + \frac{1}{Y_{21}} + \frac{1}{Y_{22}}}$$

Confidence Interval for $\ln(\phi)$

$$\ln\left(\hat{\phi}\right) \pm z_{1-\frac{\alpha}{2}} SE(\ln(\hat{\phi}))$$

Confidence Interval for ϕ

- Take exponential function of endpoints to return values to original scale.
- If CI for $\ln(\phi) = (a, b)$, then CI for ϕ is:

$$(\exp(a), \exp(b))$$

Ex. Doctor Survey

$$SE(\ln(\hat{\phi})) = \sqrt{\frac{1}{2570(0.5122)} + \frac{1}{2570(0.4878)} + \frac{1}{2645(0.5259)} + \frac{1}{2645(0.4741)}}$$

$$= 0.0399$$

Ex. Doctor Survey

- 95% CI for $\ln(\phi)$

$$\ln(0.9466) \pm 1.96(0.0399) = (-0.1331, 0.0233)$$

- 95% CI for ϕ

$$(\exp(-0.1331), \exp(0.0233)) = (0.8754, 1.0236)$$

Ex. Interpretation

- We are 95% confident the odds of returning the survey for doctors who receive a letter is between 0.8754 and 1.0236 times the odds of returning the survey for doctors who do not receive a letter in this population.
- Notice the confidence interval for the population odds ratio contains 1.
  - Implying that $p_1 = p_2$ is a reasonable conclusion given our data.