Lesson 14

Confidence Intervals of Odds Ratio and Relative Risk

Lesson 14 Outline

Lesson 14 covers

- Confidence Interval of an Odds Ratio
  - Review of Odds Ratio
  - Sampling distribution of OR on natural log scale
  - 95% Confidence Interval of OR example
- Confidence Interval of Relative Risk
  - Review of Relative Risk
  - Sampling distribution of RR on natural log scale
  - 95% Confidence Interval of RR example

Review: Odds Ratio and Relative Risk

- Both Odds Ratio and Relative Risk are measures of association (or relationship) between two nominal variables.

- The Odds Ratio is typically estimated from data collected in a Case-Control study.

- The Relative Risk can be estimated from data collected in a prospective study.

Associations between 2 nominal variable

- If there is no association between the two variables, the OR or RR = 1

- An OR or RR > 1.0 or < 1.0 indicates a *possible* statistical relationship (or association) between the two variables.

- Hypothesis tests for the OR and RR are not used to determine statistical significance of the association

- Instead, Confidence intervals of OR or RR are constructed and used to determine whether or not the association is statistically significant.

Confidence Interval of Odds Ratio

Review: Odds Ratio

■ Odds of an event =

$$\frac{\text{probability that the event occurs}}{\text{probability that the event does not occur}}$$

■ Odds Ratio

$$\frac{\text{odds of event for group 1}}{\text{odds of event for group 2}}$$

Review: Disease OR for 'Exposed' compared to 'Not Exposed'

	Disease	No Disease	Total
Exposed Group	a	b	a+b
Not Exposed Group	c	d	c+d

Odds of Disease for Exposed = a/(a+b) ÷ b/(a+b) = a/b

Odds of Disease for Not Exposed = c/(c+d) ÷ d/(c+d) = c/d

$$OR = a/b \div c/d = ad \ / \ bc$$

Confidence Interval for an Odds Ratio

- The confidence interval for an Odds Ratio has the same general formula as the Confidence Interval for a population mean or population proportion

$$\text{Point Estimate} \pm \text{Confidence Coefficient} * \text{Standard Error}$$

- The difference is that the confidence interval for the Odds Ratio is calculated on the natural log (LN) scale and then converted back to the original scale.

Sampling Distribution of the Odds Ratio

- The sampling distribution of the Odds Ratio is positively skewed.

- However, It is approximately normally distributed on the natural log scale.

- The confidence interval is calculated on the natural log scale (LN)

- After finding the limits on the LN scale, use the EXP function to find the limits on the original scale.

95% Confidence Interval of OR: the steps

1. Calculate the Odds Ratio from the data

2. Find the Natural log of the OR using the LN function in Excel

3. The confidence coefficient is from the standard normal distribution: 1.96 for a 95% confidence interval

4. Calculate the SE of LN(OR) – see next slide

5. The lower and upper limits on the log scale = LN(OR) ± 1.96* SE LN(OR) = (LL, UL)

6. Use the EXP function to find the CI limits on the original scale: EXP(LL), EXP(UL)

SE of LN(OR)

- The SE of LN(OR) is calculated from the cell counts in the 2 X 2 table:

	Cases	Controls
Exposed	a	b
Not exposed	c	d

SE of LN(OR) = $\sqrt{\dfrac{1}{a} + \dfrac{1}{b} + \dfrac{1}{c} + \dfrac{1}{d}}$

95% Confidence Interval of Odds Ratio

- First calculate the upper and lower limits on the LN scale

$$\ln(OR) \pm 1.96 * \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

- After calculating the upper and lower limits on the LN scale, use the exponential function (EXP) to find the lower and upper limits on the original scale

- If the 95% Confidence Interval does not contain the value 1.0, the association is statistically significant at alpha = 0.05

Example: Preterm delivery and Mother's SES level

■ Study background: A case-control study of the epidemiology of preterm delivery was undertaken at Yale-New Haven Hospital in Connecticut during 1977. The table on the following slide contains data for Mother's socioeconomic status (SES) for those with (cases) and without (controls) preterm delivery.

| Socio-Economic Status | Preterm Delivery | |
	Cases	Controls
Upper	11	40
Upper Middle	14	45
Middle	33	64
Lower Middle	59	91
Lower	53	58
Unknown	5	5

Construct a 95% Confidence Interval for the odds ratio for Preterm delivery comparing Lower to Upper SES

Data table for Confidence Interval of Odds Ratio of preterm delivery comparing lower to upper SES

	Preterm Delivery	
SES	Yes: Cases	No: Controls
Lower	53	58
Upper	11	40

1. Calculate the OR for Preterm Delivery: Lower SES compared to Upper SES

	Preterm Delivery	
SES	**Yes: Cases**	**No: Controls**
Lower	53	58
Upper	11	40

Odds Ratio = (53/58) / (11/40) = 3.32

2. Calculate the LN (OR) and
3. Find the confidence coefficient

2. Calculate the LN(OR)

　LN(3.32) = 1.20

　This is the point estimate for the confidence interval

3. The confidence coefficient for the LN(OR) is from the standard normal distribution.

　For a 95% confidence interval, the coefficient = 1.96

4. Calculate the SE of LN(OR)

	Preterm Delivery	
SES	Yes: Cases	No: Controls
Lower	53	58
Upper	11	40

$$\text{SE ln(OR)} = \sqrt{\frac{1}{53} + \frac{1}{58} + \frac{1}{11} + \frac{1}{40}} = 0.39$$

5. & 6. Find the Lower and Upper limits on the LN scale and the original scale

$$1.20 \pm 1.96 * 0.39 = (0.44, 1.97)$$

5. The upper and lower limits of the confidence interval on the LN scale are (0.44, 1.97)

6. Use the EXP function in Excel to find the interval limits of the 95% CI of the OR: (1.55, 7.14)

Interpretation of the CI for Preterm Delivery and SES

State the conclusion in the original scale

- The odds ratio for pre-term delivery for Lower SES compared to the Upper SES is 3.32 indicating increased odds of pre-term delivery for mothers in Lower SES

- The 95% Confidence Interval of the Odds Ratio (1.55 , 7.14) indicates that odds of pre-term delivery are significantly higher for the Lower SES group compared to the Upper SES group (at 0.05 significance level) because the CI does not contain 1.

Interpretation of 95% Confidence Interval of OR

- An Odds Ratio = 1 indicates 'no association' between the exposure and the disease.

- If the 95% confidence interval for the OR does not contain 1.0 we can conclude that there is a statistically significant* association between the exposure and the disease.

  * at the 0.05 significance level

- If the 95% confidence interval for the OR contains 1.0, the association is not significant at the 0.05 level.

Confidence Interval of Relative Risk

Review: Relative Risk

- Risk of an event =

  probability that the event occurs

- Relative Risk

  $$\frac{\text{probability that event occurs for group 1}}{\text{probability that event occurs for group 2}}$$

Review: RR Calculated from a table

	Disease	No Disease	Total
Exposed Group	a	b	a+b
Not Exposed Group	c	d	c+d

Risk of Disease for Exposed = a/(a+b)

Risk of Disease for Not exposed = c/(c+d)

RR = a/(a+b) ÷ c/(c+d)

Confidence Interval for Relative Risk

- Like the Odds Ratio, the sampling distribution of the Relative Risk is positively skewed but is approximately normally distributed on the natural log scale.

- Constructing a Confidence Interval for the Relative Risk is similar to constructing a CI for the Odds Ratio except that there is a different formula for the SE.

95% Confidence Interval of RR

1. Estimate the RR from the data

2. Find the natural log of RR: LN(RR)

3. The confidence coefficient is from the standard normal distribution: 1.96 for a 95% confidence interval

4. Calculate the SE of LN(RR) – see next slide

5. Calculate the lower and upper limits on the LN scale: LN(RR) ± 1.96* SE LN(RR)

6. Use the EXP function to find the limits on the original scale: EXP(LL), EXP (UL)

SE of LN(RR)

■ The SE of LN(RR) is calculated from the cell counts in the 2 X 2 table:

	Disease	**No Disease**	**Total**
Exposed Group	a	b	a+b
Not Exposed Group	c	d	c+d

$$SE \ln(RR) = \sqrt{\frac{b}{a(a+b)} + \frac{d}{c(c+d)}}$$

This formula is algebraically equivalent to the formula in text on pg. 201

95% Confidence Interval of Relative Risk

- First calculate the upper and lower limits on the LN scale

$$\ln(RR) \pm 1.96 * \sqrt{\frac{b}{a(a+b)} + \frac{d}{c(c+d)}}$$

- After calculating the upper and lower limits on the LN scale, use the exponential function to find the limits on the original scale

- If the 95% Confidence Interval does not contain the value 1.0, the association is statistically significant at alpha = 0.05

Example: 95% CI for RR of Myocardial Infarction (MI)

- Background: Physicians enrolled in the Physician's Health Study were randomly assigned to take a daily aspirin or placebo. The table provides the number with MI in each group.

Group	Myocardial Infarction (MI)		
	Yes	No	
Aspirin	139	10898	11037
Placebo	239	10795	11034
	378	21693	22071

- Calculate the RR and 95% CI for the RR for MI (aspirin group compared to placebo group)

1. Calculate the RR for MI: aspirin group compared to placebo

Group	Myocardial Infarction (MI)		
	Yes	No	
Aspirin	139	10898	11037
Placebo	239	10795	11034
	378	21693	22071

RR = (139/11037) / (239/11034) = 0.581

2. Calculate the LN (RR) and
3. Find the confidence coefficient

2. Calculate the LN(RR)

   LN(0.581) = -0.543

   This is the point estimate for the confidence interval

3. The confidence coefficient for the LN(RR) is from the standard normal distribution.

   For a 95% confidence interval, the coefficient = 1.96

4. Calculate the SE LN(RR)

Group	Myocardial Infarction (MI)		
	Yes	No	
Aspirin	139	10898	11037
Placebo	239	10795	11034
	378	21693	22071

$$\text{SE ln(RR)} = \sqrt{\frac{10898}{139 * 11037} + \frac{10795}{239 * 11034}} = 0.106$$

5. & 6. Find the Lower and Upper limits on the LN scale and the original scale

$$-0.543 \pm 1.96 * 0.106 = (-0.75, -0.336)$$

5. The upper and lower limits of the confidence interval on the LN scale are

(-0.750, -0.336)

6. Use the EXP function in Excel to find the interval limits of the 95% CI of the RR: (0.472, 0.715)

Interpretation of 95% CI of RR for MI

- The Relative Risk estimate = 0.58 which indicates that physicians in the aspirin group had a lower risk of MI than physicians in the placebo group.

- The 95% confidence interval indicates that the decreased risk related to daily aspirin use is significant (at alpha = 0.05) since the interval (0.472, 0.715) does not contain 1.

General Interpretation of 95% Confidence Interval of RR

- A Relative Risk = 1.0 indicates 'no association' between the exposure and the disease.

- If the 95% confidence interval for the RR does not contain 1.0 we can conclude that there is a statistically significant* association between the exposure and the disease.

  * at the 0.05 significance level

- If the 95% confidence interval for the RR contains 1.0, the association is not significant at the 0.05 level.

Notes on the OR and RR CI:

- The RR estimate will not be exactly in the middle of the confidence interval on the original scale
  - The LN(RR) is exactly in the middle of the interval on the LN scale
- The OR estimate will not be exactly in the middle of the confidence interval on the original scale
  - The LN(OR) is exactly in the middle of the interval on the LN scale
- Check that the RR or OR estimate is contained in the final interval – it's easy to forget the EXP step

Readings and Assignments

- Reading: Chapter 8 pgs 200-201
- Work through the Lesson 14 practice exercises
- Excel Module 14 works through the examples in this lesson for the confidence intervals of OR and RR
  - Excel functions: LN and EXP
  - There is no CI function for the RR and OR – you just work through the steps
- Complete Homework 10