Relative Measures

Department of Statistics, University of South Carolina

Stat 705: Data Analysis II

Outline

- Relative measures
- The relative risk
- The odds ratio

Motivation

- Consider a randomized trial where 40 subjects were randomized (20 each) to two drugs with the same active ingredients but different excipients
- Consider counting the number of participants with side effects for each drug

	Side Effects	None	Total
Drug A	11	9	20
Drug B	5	15	20
Total	16	24	40

Comparing two binomials

- Consider now testing whether the proportion of side effects is the same in the two groups.
- Let $X \sim Binomial(n_1, p_1)$ and $\widehat{p_1} = \frac{X}{n_1}$
- Let $Y \sim Binomial(n_2, p_2)$ and $\widehat{p_2} = \frac{Y}{n_2}$

$n_{11} = X$	$n_{12} = n_1 - X$	$n_1 = n_{1+}$
$n_{21} = Y$	$n_{22} = n_2 - Y$	$n_2 = n_{2+}$
n_{+1}	n_{+2}	

The relative risk

- Last time, we considered the absolute change in the proportions, what about relative changes?
- Relative changes are often of more interest than absolute. For example when both proportions are small
- The **relative risk** is defined as $\frac{p_1}{p_2}$
- The natural estimator for the relative risk is

$$\widehat{RR} = \frac{\hat{p}_1}{\hat{p}_2} = \frac{X/n_1}{Y/n_2}$$

- The standard error for $\log \widehat{RR}$ is

$$\widehat{SE}_{\log \widehat{RR}} = \left(\frac{(1-p_1)}{p_1 n_1} + \frac{(1-p_2)}{p_2 n_2}\right)^{\frac{1}{2}}$$

- Exponentiate the resulting interval to get an interval for the RR

The odds ratio

- The odds in favor of an event are the probability that the event will happen to the probability that it will not happen.

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

- The odds ratio is defined as

$$\frac{\text{odds of SE Drug A}}{\text{odds of SE Drug B}} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{p_1(1-p_2)}{p_2(1-p_1)}$$

  SE: side effect

- The sample odds ratio simply plugs in the estimates for p_1 and p_2. This work out to have a convenient form (cross product ratio).

$$\widehat{OR} = \frac{\widehat{p}_1(1-\widehat{p}_2)}{\widehat{p}_2(1-\widehat{p}_1)} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

- The standard error of $\log \widehat{OR}$ is

$$\widehat{SE_{\log \hat{OR}}} = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

- Exponentiate the resulting interval to obtain an interval for the OR

Some Comments

- Notice that the sample and true odds ratios do not change if we transpose the rows and columns
- For both the OR and the RR, taking logs helps with adherence to the error rate
- The intervals for the log RR and log OR can be obtained by taking
$$Estimate \pm Z_{1-\alpha/2} SE_{Estimate}$$
- Exponentiating yields an interval for the OR or RR.
- Though logging helps, there intervals still don't perform very well

Example-RR

- For the relative risk, $\widehat{p_A} = \frac{11}{20} = 0.55$, $\widehat{p_B} = \frac{5}{20} = 0.25$
- $\widehat{RR}_{A/B} = 0.55/0.25 = 2.2$
- $\widehat{SE_{\log \hat{RR}_{A/B}}} = \sqrt{\frac{1-0.55}{0.55\times 20} + \frac{1-0.25}{0.25\times 20}} = 0.44$
- Interval foe the log RR :

$$\log 2.2 \pm 1.96 \times 0.44 = (-0.07, 1.65)$$

- Interval for the RR: (0.93, 5.21)

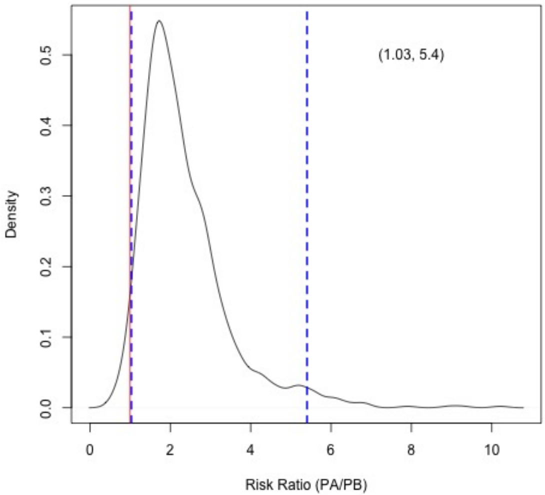
Example-OR

- $\widehat{OR_{A/B}} = \frac{11 \times 15}{9 \times 5} = 3.67$
- $SE_{\widehat{\log \hat{OR}_{A/B}}} = \sqrt{\frac{1}{11} + \frac{1}{9} + \frac{1}{5} + \frac{1}{15}} = 0.68$
- Interval for log OR : $\log 3.67 \pm 1.96 \times 0.68 = (-0.04, 2.64)$
- Interval for OR: (0.96, 14.01)

R Code

```
x <- 11; n1 <- 20; alpha1 <- 1; beta1 <- 1
y <- 5; n2 <- 20; alpha2 <- 1; beta2 <- 1
p1 <- rbeta(1000, x + alpha1, n1 - x + beta1)
p2 <- rbeta(1000, y + alpha2, n2 - y + beta2)
rd <- p2 - p1
plot(density(rd))
quantile(rd, c(.025, .975))
mean(rd)
median(rd)
```

Bayesian Posterior Interval for RR



Bayesian Posterior Interval for OR