Spring 2014
Stat 8004: Statistical Methods II
Lecture 8

1 Stratified-Adjusted Analysis

1.1 Introduction

It is said that *randomization* is a core principle in statitsical theory. By randomization, one can systematically eliminate all the imbalances in the data. However, sometimes, it is hard or impractical to randomize.

Example: New operation on knee injury. Dr. Wright would like to study the successful rates of different operations on the knee injury. Patients with knee injuries consult him, but decide the treatment by themselves. Dr. Wright believed that the new operation has a higher successful rate. However, after he collected the data, he found that the traditional one has a higher estimated successful rate. As a statistician, what would be your suggestion?

Note that in the previous example, no *randomization* is conducted. The patients chose their own treatment. It turned out that the patients with direct injuries prefer the new operation. Since their injuries are more serious, the successful rate is generally low for any treatment. On the other hand, the patients with twist injuries prefer the traditional operation. Since most of these injuries are minor, the successful rate is generally high for any treatment. In other words, the *imblances* in Dr. Wright's data set causes some problems. One solution to solve these problems is to use *stratified studies*.

1.2 Marginal Unadjusted Analysis

In the knee injury study, Dr. Wright collected the data in Table 1. If we simply conbine these two tables together, the data is as in Table 2. If we performed strata-specific analysis

Direct Injuries

Exposure	Response: Success	Partially Success	
New	40	30	70
Old	15	15	30
	55	45	100

Twist Injuries

Exposure	Response: Success	Partially Success	
New	15	5	20
Old	55	25	80
	70	30	100

Table 1: Stratiefied Data of Knee Injuries and Operations

All Injuries

Exposure	Response: Success	Partially Success	
New	55	35	90
Old	70	40	110
	125	75	200

Table 2: Marginal Data of Knee Injuries and Operations

as well as marginal analysis, we have the following results in Table 3.

Marginally, the differences between treatment groups is not statistically significant. It appears that marginally, the traditional operation has a slightly higher successful rate that the new operation, even though within each stratum, the new operation seems to be slightly better. The marginal analysis ignores the imbalance between treatment groups in the nubmers of subjects. One might ask whether the nature or significance of the treatment group effect is altered in any way after adjusting for these imblances within strata.

1.3 Stratified-Adjusted Tests

1.3.1 Mantel-Haenszel Test

Table 4 is a general form of a 2×2 table of a stratified study. In the knee injury example, there are two strata, the first one includes the patient with direct injuries, and the second includes the patients with twist injuries.

Within the jth strata, we can do inferences as abover. However, analysis within-strata lead to multiple tests of siginificance. It doesn't help to answer the general question which

Measure	Stratum 1	2	Marginal
Risk difference $\widehat{RD}$	0.07	0.06	-0.03
95% CI for RD	-0.14,0.28	-0.15,0.28	-0.16,0.11
Relative risk $\widehat{RR}$	1.14	1.09	0.96
95% CI for RR	0.75,1.72	0.81,1.46	0.77,1.19
Odds ratio $\widehat{OR}$	1.33	1.36	0.89
95% CI for OR	0.57,3.14	0.45,4.17	0.50,1.6

Table 3: Measurement of Association Within Each Stratum and in the Marginal Unadjusted Analysis

Exposure	Response +	−	
1	a_j	b_j	n_{1j}
2	c_j	d_j	n_{2j}
	m_{1j}	m_{2j}	N_j

Table 4: the jth strata of a stratified study

treatment is better.

The Mantael-Haenszel test is a test of the global null hypothesis:

$$\mathrm{H}_0 : \pi_{1j} = \pi_{2j} \ (OR_j = 1), \quad \text{for all } j = 1, \ldots, K.$$

In the knee injury example, $K = 2$.

As shown in Lecture 8, when $\varphi_j = OR_j = 1$, within the jth stratum,

$$\begin{aligned} E_j &= \mathbb{E}\left(a_j\right) = \frac{n_{1j} m_{1j}}{N_j} \\ \hat{V}_{cj} &= \widehat{\mathrm{Var}}(a_j) = \frac{n_{1j} n_{2j} m_{1j} m_{2j}}{N_j^2 (N_j - 1)} \end{aligned}$$

Within the jth stratu, under H_0, asymptotically for large N_j and for fixed K,

$$a_j - E_j \overset{d}{\approx} N(0, \hat{V}_{cj})$$

Since the strata are independent,

$$\sum_j (a_j - E_j) \overset{d}{\approx} N\left(0, \sum_j \hat{V}_{cj}\right)$$

Therefore, the stratified-adjusted Mantel-Haenszel test, conditional on n_{1j}, n_{2j}, m_{1j} fixed, is

$$X^2_{C(MH)} = \frac{\left[\sum_j (a_j - E_j)\right]^2}{\sum_j \hat{V}_{cj}} = \frac{\left[\sum_j (a_j - n_{1j} m_{1j})/N_j\right]^2}{\sum_j \left(\frac{n_{1j} n_{2j} m_{1j} m_{2j}}{N_j^2 (N_j - 1)}\right)} = \frac{[a_+ - E_+]^2}{\hat{V}_{c+}},$$

where $a_+ = \sum_j a_j$, $E_+ = \sum_j E_j$ and $\hat{V}_{c+} = \sum_j \hat{V}_{cj}$. Asymptotically, under H_0,

$$X^2_{C(MH)} \overset{d}{\approx} \chi^2(1).$$

Example: New Operation on Knee Injuries.

```
Mantel-Haenszel chi-squared test without continuity correction

data:  knee
Mantel-Haenszel X-squared = 0.723, df = 1, p-value = 0.3951
alternative hypothesis: true common odds ratio is not equal to
  1
95 percent confidence interval:
 0.6810536 2.6555345
sample estimates:
common odds ratio
         1.344828
```

1.3.2 Cochran's Test

As shown in Lecture 8, we can also use Cochran's test to test $\mathrm{H}_0 : OR = 1$.

Within the jth stratum, under product binomial distribution,

$$\hat{E}_j = \widehat{\mathbb{E}}(a_j) = n_{1j}m_{1j}/N_j$$
$$\hat{V}_{uj} = \widehat{\mathrm{Var}}(a_j) = \frac{n_{1j}n_{2j}m_{1j}m_{2j}}{N_j^3}$$

Similarly, we aggregated over strata, the stratified-adjusted Cochran's test is

$$X^2_{U(C)} = \frac{\left[\sum_j (a_j - \hat{E}_j)\right]^2}{\sum_j \hat{V}_{uj}} = \frac{\left[\sum_j (a_j - n_{1j}m_{1j})/N_j\right]^2}{\sum_j \left(\frac{n_{1j}n_{2j}m_{1j}m_{2j}}{N_j^3}\right)} = \frac{[a_+ - \hat{E}_+]^2}{\hat{V}_{u+}},$$

where $a_+ = \sum_j a_j$, $\hat{E}_+ = \sum_j \hat{E}_j$ and $V_{u+} = \sum_j V_{uj}$. Asymptotically, under H_0,

$$X^2_{U(C)} \overset{d}{\approx} \chi^2(1).$$

The only difference between $X^2_{C(MH)}$ and $X^2_{U(C)}$ is in the denominators. Since

$$\hat{V}_{uj} = \frac{\hat{V}_{cj}(N_j - 1)}{N_j},$$

the two tests are asymptotically equivalent. They are often referred to interchangeably as the *Cochran-Mantel-Haenszel-Test*.

1.4 Stratified Adjusted Estimators

1.4.1 Mantel-Haenszel Estimators

Table 3 gives out $\widehat{RD}$, $\widehat{RR}$ and $\widehat{OR}$ for each stratum and for marginal analysis. The marginal estimators doesn't adjust for imbalances between groups in number of subjects. Mantael and Haenzsel (1959) prensented a "heuristic" estimators of the assumed common odds ratio:

$$\widehat{OR}_{MH} = \frac{\sum_j a_j d_j / N_j}{\sum_j b_j c_j / N_j}$$

Note that

$$\widehat{OR}_j = \frac{a_j d_j}{b_j c_j}$$

Define

$$\hat{v}_j = \frac{b_j c_j / N_j}{\sum_l b_l c_l / N_l},$$

which can be viewed as weights since $\sum_j \hat{v}_j = 1$.

It turns out that

$$\widehat{OR}_{MH} = \sum_j \hat{v}_j \cdot \widehat{OR}_j,$$

which is a weighted sum of $\widehat{OR}_j$.

Likewise, when a constant relative risk is assumed, the Mantel-Haenszel estimate of the common relative risk is

$$\widehat{RR}_{MH} = \frac{\sum_j a_j n_{2j} / N_j}{\sum_j b_j n_{1j} / N_j} = \sum_j \hat{v}_j \cdot \widehat{RR}_j,$$

where

$$\hat{v}_j = \frac{b_j n_{1j} / N_j}{\sum_l b_l n_{1l} / N_l}$$
$$\widehat{RR}_j = \frac{a_j n_{2j}}{b_j n_{1j}}$$

1.4.2 Large Sample Variance of Log Odds Ratio

Let $OR_{MH} = \mathbb{E}\,(\widehat{OR}_{MH})$, and $\theta = \log(OR_{MH})$. Also let $\theta_j = \log(OR_j)$, where we know that $OR_j \approx \mathbb{E}\,(\widehat{OR}_j)$.

Using delta method, we can show that

$$\text{Var}(\widehat{OR}_{MH}) \approx (OR_{MH})^2 \text{Var}(\hat{\theta}),$$

which can be estimated by $\widehat{OR}_{MH}^2 \widehat{\text{Var}}(\hat{\theta})$.

Note that

$$\widehat{OR}_{MH} = \sum_j \hat{v}_j \cdot \widehat{OR}_j$$

If we treat $\hat{v}_j$ as known (fixed, not random variable), then the asymptotic variance

$$\begin{aligned}
\text{Var}(\widehat{OR}_{MH}) &= \sum_j \hat{v}_j^2 \cdot \text{Var}(\widehat{OR}_j) \\
&\approx \sum_j \hat{v}_j^2 \cdot (OR_j)^2 \cdot \text{Var}(\log(\widehat{OR}_j)) \\
&= \sum_j \hat{v}_j^2 \cdot (OR_j)^2 \cdot \text{Var}(\hat{\theta}_j)
\end{aligned}$$

Therefore,

$$\text{Var}(\hat{\theta}) \approx \frac{\text{Var}(\widehat{OR}_{MH})}{(\widehat{OR}_{MH})^2},$$

which can be estimated as $\widehat{\text{Var}}(\widehat{OR}_{MH})/(\widehat{OR}_{MH})^2$.

Then based on these results, we can construct confidence intervals for OR_{MH}. The $(1-\alpha)$ confidence limits for $\hat{\theta}$ is

$$\hat{\theta}_l = \hat{\theta} - z_{1-\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\theta})}$$
$$\hat{\theta}_u = \hat{\theta} + z_{1-\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\theta})}$$

And it can be transformed back to the confidence limits for OR_{MH}: $(\exp(\hat{\theta}_l),\ \exp(\hat{\theta}_u))$.

Association Measure	Stratum 1	Stratum 2	Mantel-Haenszel Estimate	95% C.I.
Odds Ratio	1.33	1.36	1.34	0.68,2.66
$\hat{v}_j$	0.62	0.38		
$\widehat{Var}(\widehat{OR}_{MH})$			0.22	

Table 5: Odds Ratios Within Strata and the Mantel-Haenszel Adjusted Estimates

1.5 Nature of Covariate Adjustment

1.5.1 Confounding and Effect Modification

In the knee injury example, the stratification variable is the type of injuries (direct injury and twist injury). Within each stratum, the new operation has a sligtly higher succefully rate. But if you simply combine the data and do not adjust for the imbanlances in the number of subjects, you will find that the old treatment has a slightly higher successful rate. This phenomenon is called the *Simpson's paradox*.

Confoundings:

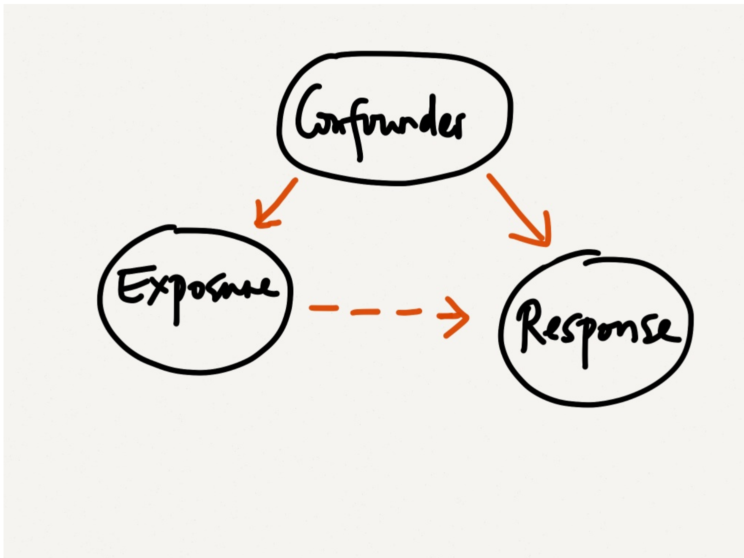

Figure 1: Confounding

Effect Modifications: OR_j are different across stratum. Then we say stratification variable modifies or interacts with the outcome.