# STAT 226 Lecture 25

## Section 8.1 Comparing Dependent Proportions

Yibi Huang

# Example (Two Case Control Studies on Hodgkin's Disease)

Study 1 by Vianna et al. (1971) :

| | Tonsillectomy | | |
|---|---|---|---|
| | Yes | No | Total |
| Cases | 67 | 34 | 101 |
| Controls | 43 | 64 | 109 |

$$\widehat{\theta} = \frac{67 \times 64}{34 \times 43} \approx 2.93$$

$$\text{Pearson's } X^2 = 13.23$$

$$df = 1$$

$$P\text{-value} = 0.00028$$

Study 2 by Johnson & Johnson (1972):

| | Tonsillectomy | | |
|---|---|---|---|
| | Yes | No | Total |
| Cases | 41 | 44 | 85 |
| Controls | 33 | 52 | 85 |

$$\widehat{\theta} = \frac{41 \times 52}{44 \times 33} \approx 1.47$$

$$\text{Pearson's } X^2 = 1.17$$

$$df = 1$$

$$P\text{-value} = 0.279$$

Why did the 2 studies reach inconsistent conclusions?

- In study 1, controls were **matched as a group** according to age, sex, race, residence, date of hospitalization, and no cancer
- In study 2, controls were patient's same sex sibling closest in age. To reflect dependence, data in Study 2 should be displayed as

| | | Sibling had tonsillectomy? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Hodgkin's patient had tonsillectomy? | Yes | 26 | 15 | 41 |
| | No | 7 | 37 | 44 |
| | Total | 33 | 52 | 85 |

  This table shows high correlation of tonsillectomy status within a sibling pair. Most pairs of siblings had the same tonsillectomy status.
- concordant pairs tell us nothing about the relationship between the disease and the risk factor.
  Sufficient to examine discordant pairs only

Population probabilities:

| | | Sibling had tonsillectomy? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Hodgkin's patient had tonsillectomy? | Yes | $\pi_{11}$ | $\pi_{12}$ | $\pi_{1+}$ |
| | No | $\pi_{21}$ | $\pi_{22}$ | $\pi_{2+}$ |
| | Total | $\pi_{+1}$ | $\pi_{+2}$ | 1 |

Discussion

- What was the goal of the study?
- Which two proportions were we interested in comparing?

Population probabilities:

| | | Sibling had tonsillectomy? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Hodgkin's patient had tonsillectomy? | Yes | $\pi_{11}$ | $\pi_{12}$ | $\pi_{1+}$ |
| | No | $\pi_{21}$ | $\pi_{22}$ | $\pi_{2+}$ |
| | Total | $\pi_{+1}$ | $\pi_{+2}$ | 1 |

Discussion

- What was the goal of the study?
- Which two proportions were we interested in comparing?

Compare dependent samples by making inference about $\pi_{1+} - \pi_{+1}$. There is *marginal homogeneity* if $\pi_{1+} = \pi_{+1}$.

# McNemar's Test

Under $H_0$: marginal homogeneity,

$$\pi_{1+} = \pi_{+1} \iff \pi_{12} = \pi_{21} \iff \frac{\pi_{12}}{\pi_{12} + \pi_{21}} = \frac{1}{2}$$

Under $H_0$, each of $n^* = n_{12} + n_{21}$ observations has probability $1/2$ of contributing to $n_{12}$ and $1/2$ of contributing to $n_{21}$:

$$n_{12} \sim \text{Binomial}\left(n^*, \frac{1}{2}\right),\ \text{mean} = \frac{n^*}{2},\ \text{std dev} = \sqrt{n^*\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}$$

By normal approx. to binomial, for large $n^*$,

$$z = \frac{n_{12} - n^*/2}{\sqrt{n^*\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}} = \frac{n_{12} - n_{21}}{\sqrt{n_{12} + n_{21}}} \sim N(0, 1)$$

or equivalently

$$z^2 = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} \sim \chi^2_1$$

# Example (Two Case Control Studies on Hodgkin's Disease)

| | | Sibling had tonsillectomy? Yes | No | Total |
|---|---|---|---|---|
| Hodgkin's patient had tonsillectomy? | Yes | 26 | 15 | 41 |
| | No | 7 | 37 | 44 |
| | Total | 33 | 52 | 85 |

For testing H$_0$: $\pi_{1+} = \pi_{+1}$ v.s. H$_a$: $\pi_{1+} \neq \pi_{+1}$

$$z = \frac{n_{12} - n_{21}}{\sqrt{n_{12} + n_{21}}} = \frac{15 - 7}{\sqrt{15 + 7}} = 1.706, \quad (z^2 = 2.909,\ df = 1)$$

$$P\text{-value} = 0.088.$$

Though still insignificant, Study 2 is more consistent w/ Study 1.

# McNemar's Test in R

```
hodgkin1972 = matrix(c(26,7,15,37), nrow=2)
hodgkin1972
     [,1] [,2]
[1,]   26   15
[2,]    7   37
mcnemar.test(hodgkin1972, correct = FALSE)

    McNemar's Chi-squared test

data:  hodgkin1972
McNemar's chi-squared = 2.9091, df = 1, p-value = 0.08808
```

# Exact McNemar's Test

When $n^* = n_{12} + n_{21}$ are small, it's more reliable to use exact binomial tests since under $H_0$: $\pi_{12} = \pi_{21}$, the exact distribution of $n_{12}$ is

$$n_{12} \sim \text{Binomial}\left(n^*, \frac{1}{2}\right)$$

For the Hodgkin's Disease study, the exact two-sided $P$-value is

```
binom.test(15, 22, p=0.5)

    Exact binomial test

data:  15 and 22
number of successes = 15, number of trials = 22, p-value = 0.1338
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.4512756 0.8613535
sample estimates:
probability of success
             0.6818182
```

# Exact McNemar's Test

$P$-values given by McNemar's test with continuity correction are closer to the $P$-values given by exact McNemar's test.

```
mcnemar.test(hodgkin1972, correct = TRUE)

    McNemar's Chi-squared test with continuity correction

data:  hodgkin1972
McNemar's chi-squared = 2.2273, df = 1, p-value = 0.1356
```

# CI for $\pi_{1+} - \pi_{+1}$

Estimate $\pi_{1+} - \pi_{+1}$ by diff. of sample proportions, $\widehat{\pi}_{1+} - \widehat{\pi}_{+1}$.

$$\widehat{\pi}_{1+} - \widehat{\pi}_{+1} = \frac{n_{1+}}{n} - \frac{n_{+1}}{n} = \frac{n_{12} - n_{21}}{n}$$

$$\text{SE} = \frac{1}{n}\sqrt{n_{12} + n_{21} - \frac{(n_{12} - n_{21})^2}{n}}$$

**Example (Hodgkin's Disease)**

| | | |
|---|---|---|
| $n_{11}$ | $n_{12}$ | |
| $n_{21}$ | $n_{22}$ | |
| | | $n$ |

=

| | | |
|---|---|---|
| 26 | 15 | |
| 7 | 37 | |
| | | 85 |

$$\widehat{\pi}_{1+} - \widehat{\pi}_{+1} = \frac{15 - 7}{85} \approx 0.094$$

$$\text{SE} = \frac{1}{85}\sqrt{15 + 7 - \frac{(15 - 7)^2}{85}} \approx 0.054$$

95% CI: $0.094 \pm (1.96)(0.054) = (-0.0118, 0.1998)$

# How is the SE of the Matched-Pairs CI Derived?

$$(n_{11}, n_{12}, n_{21}, n_{22}) \sim \text{Multinomial}\Big(n,\ (\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22})\Big)$$

$$\Longrightarrow \begin{cases} \mathrm{Var}(n_{ij}) = n\pi_{ij}(1-\pi_{ij}) \\ \mathrm{Cov}(n_{ij}, n_{i'j'}) = -n\pi_{ij}\pi_{i'j'} \quad \text{if } i \neq i' \text{ or } j \neq j' \end{cases}$$

$$\begin{aligned}
\mathrm{Var}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1}) &= \mathrm{Var}\left(\frac{n_{12} - n_{21}}{n}\right) = \frac{\mathrm{Var}(n_{12} - n_{21})}{n^2} \\
&= \frac{\mathrm{Var}(n_{12}) + \mathrm{Var}(n_{21}) - 2\,\mathrm{Cov}(n_{12}, n_{21})}{n^2} \\
&= \frac{n\pi_{12}(1-\pi_{12}) + n\pi_{21}(1-\pi_{21}) + 2n\pi_{12}\pi_{21}}{n^2} \\
&= \frac{\pi_{12} + \pi_{21} - (\pi_{12}^2 - 2\pi_{12}\pi_{21} + \pi_{21}^2)}{n} \\
&= \frac{\pi_{12} + \pi_{21} - (\pi_{12} - \pi_{21})^2}{n} \quad \text{(ctd next slide)}
\end{aligned}$$

# How is the SE of the Matched-Pairs CI Derived? (Cont'd)

$$
\begin{aligned}
\mathrm{Var}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1}) &= \frac{\pi_{12} + \pi_{21} - (\pi_{12} - \pi_{21})^2}{n} \\
\widehat{\mathrm{Var}}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1}) &= \frac{\widehat{\pi}_{12} + \widehat{\pi}_{21} - (\widehat{\pi}_{12} - \widehat{\pi}_{21})^2}{n} \\
&= \frac{\frac{n_{12}}{n} + \frac{n_{21}}{n} - (\frac{n_{12}}{n} - \frac{n_{21}}{n})^2}{n} \\
&= \frac{\frac{n_{12}}{n} + \frac{n_{21}}{n} - \frac{(n_{12}-n_{21})^2}{n^2}}{n} \times \frac{n}{n} \\
&= \frac{n_{12} + n_{21} - \frac{(n_{12}-n_{21})^2}{n}}{n^2}
\end{aligned}
$$

So

$$
\mathrm{SE} = \sqrt{\mathrm{Var}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1})} = \frac{1}{n}\sqrt{n_{12} + n_{21} - \frac{(n_{12} - n_{21})^2}{n}}
$$

# Two-Sample Designs v.s Matched-Pair Designs

$$\mathrm{Var}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1}) = \mathrm{Var}(\widehat{\pi}_{1+}) + \mathrm{Var}(\widehat{\pi}_{+1}) - 2\mathrm{Cov}(\widehat{\pi}_{1+}, \widehat{\pi}_{+1})$$

where $\mathrm{Var}(\widehat{\pi}_{1+}) = \dfrac{\pi_{1+}(1-\pi_{1+})}{n}$, $\mathrm{Var}(\widehat{\pi}_{+1}) = \dfrac{\pi_{+1}(1-\pi_{+1})}{n}$,

$$\begin{aligned}
\mathrm{Cov}(\widehat{\pi}_{1+}, \widehat{\pi}_{+1}) &= \mathrm{Cov}\left(\frac{n_{1+}}{n}, \frac{n_{+1}}{n}\right) = \mathrm{Cov}\left(\frac{n_{11}+n_{12}}{n}, \frac{n_{11}+n_{21}}{n}\right) \\
&= \frac{1}{n^2}\mathrm{Cov}\left(n_{11}+n_{12}, n_{11}+n_{21}\right) \\
&= \frac{1}{n^2}\left[\mathrm{Var}(n_{11}) + \mathrm{Cov}(n_{11}, n_{21}) + \mathrm{Cov}(n_{12}, n_{11}) + \mathrm{Cov}(n_{12}, n_{21})\right] \\
&= \frac{1}{n^2}\left[n\pi_{11}(1-\pi_{11}) - n\pi_{11}\pi_{21} - n\pi_{12}\pi_{11} - n\pi_{12}\pi_{21}\right] \\
&= \frac{1}{n}[\pi_{11}\underbrace{(1-\pi_{11}-\pi_{12}-\pi_{21})}_{\pi_{22}} - \pi_{12}\pi_{21}] \\
&= \frac{\pi_{11}\pi_{22} - \pi_{12}\pi_{21}}{n}
\end{aligned}$$

# Two-Sample Designs v.s Matched-Pair Designs (Cont'd)

Often matched-pairs exhibit positive association (odds-ratio greater than 1), i.e., $\pi_{11}\pi_{22} > \pi_{12}\pi_{21}$, so covariance term is negative.

$$
\begin{aligned}
& \mathrm{Var}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1}) \quad \text{for matched-pairs design} \\
={}& \frac{1}{n}[\pi_{1+}(1-\pi_{1+}) + \pi_{+1}(1-\pi_{+1}) - 2(\underbrace{\pi_{11}\pi_{22} - \pi_{12}\pi_{21}}_{\textit{usually} >0})] \\
\leq{}& \frac{1}{n}\left[\pi_{1+}(1-\pi_{1+}) + \pi_{+1}(1-\pi_{+1})\right] \\
={}& \mathrm{Var}(\widehat{\pi}_{1+} - \widehat{\pi}_{+1}) \quad \text{for two indep. samples of size } n \text{ each}
\end{aligned}
$$

If matched-pairs exhibit positive association, matched-pairs designs are more efficient than 2-sample designs.

# Example (Opinions Relating to Environment)

In 2000 General Social Survey, subjects were asked whether, to help the environment, they would be willing to (1) pay higher taxes or (2) accept a cut in living standards.

| | Yes | No | Total | % Yes |
|---|---|---|---|---|
| Pay Higher Taxes? | 334 | 810 | 1144 | 29.2% |
| Cut Living Standards? | 359 | 785 | 1144 | 31.4% |

- Q: Which option were there more people willing to accept?
- The two sample percentages are *dependent* because the same subjects were asked both questions. There were 1144 subjects only, not $1144 + 1144$.

To reflect dependence, data should be displayed as

| | | Willing to Cut Living Standards? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Willing to Pay Higher Taxes? | Yes | 227 | 132 | 359 |
| | No | 107 | 678 | 785 |
| | Total | 334 | 810 | 1144 |

# Example (Opinions Relating to Environment)

| | | Willing to Cut Living Standards? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Willing to Pay Higher Taxes? | Yes | 227 | 132 | 359 |
| | No | 107 | 678 | 785 |
| | Total | 334 | 810 | 1144 |

Estimates for $\pi_{1+}$ = proportion willing to pay higher taxes, and $\pi_{+1}$ = proportion willing to cut living standards, are respectively

$$\widehat{\pi}_{1+} = \frac{359}{1144} \approx 0.314, \quad \text{and } \widehat{\pi}_{+1} = \frac{334}{1144} \approx 0.292.$$

The 95% confidence interval for $\pi_{1+} - \pi_{+1}$ is NOT

$$(0.314-0.292) \pm 1.96\sqrt{\frac{0.314(1-0.314)}{1144} + \frac{0.292(1-0.292)}{1144}} \approx 0.022 \pm 0.038$$

The correct CI is

$$(0.314 - 0.292) \pm 1.96\frac{1}{1144}\sqrt{132 + 107 - \frac{(132-107)^2}{1144}} \approx 0.022 \pm 0.026.$$

# Example (Opinions Relating to Environment)

**Table I**

| | Yes | No | Total |
|---|---|---|---|
| *Pay Higher Taxes?* | 334 | 810 | 1144 |
| *Cut Living Standards?* | 359 | 785 | 1144 |

**Table II**

| *Pay Higher Taxes?* | *Cut Living Standards?* Yes | No | Total |
|---|---|---|---|
| Yes | 227 | 132 | 359 |
| No | 107 | 678 | 785 |
| Total | 334 | 810 | 1144 |

To test whether there were more people willing to pay higher taxes or more people willing to cut living standards, we should . . .

a. perform Pearson's $X^2$ test on Table I
b. perform Pearson's $X^2$ test on Table II
c. perform McNemar's test on Table I
d. perform McNemar's test on Table II

Which one is correct?

# Example (Opinions Relating to Environment)

**Table I**

| | Yes | No | Total |
|---|---|---|---|
| *Pay Higher Taxes?* | 334 | 810 | 1144 |
| *Cut Living Standards?* | 359 | 785 | 1144 |

**Table II**

| *Pay Higher Taxes?* | *Cut Living Standards?* Yes | No | Total |
|---|---|---|---|
| Yes | 227 | 132 | 359 |
| No | 107 | 678 | 785 |
| Total | 334 | 810 | 1144 |

To test whether those willing to pay higher taxes were more or less willing to cut living standards than those not willing to pay higher taxes, we should . . .

a. perform Pearson's $X^2$ test on Table I
b. perform Pearson's $X^2$ test on Table II
c. perform McNemar's test on Table I
d. perform McNemar's test on Table II

Which one is correct?