Chap. 8 Models for Matched Pairs

ex) Crossover study to compare drug with placebo. 86 subjects randomly assigned to receive drug then placebo or else placebo then drug Binary response (S, F) for each

Treatment	S	F	Total
Drug	61	25	86
Placebo	22	64	86

Methods so far (eg., X^2 , G^2 test of indep., C.I. for θ , logistic regression) assume independent samples, they are inappropriate for dependent samples (eg., sample subjects in each sample, which yield matched pairs)

To reflect dependence, display data as 86 observations rather than 2×86 observations.

Placebo
$$S$$
 F

Drug
 S
 12
 49
 61 (71%)
 F
 10
 15
 25
 22
 64
 86
 $(*)$

Population probabilities

Placebo
$$S$$
 F T_{11} T_{12} T_{1+} T_{21} T_{22} T_{2+} T_{21} T_{22} T_{2+} T_{21} T_{22} T_{22} T_{22} T_{22} T_{23}

Compare dependent samples by making inference about

$$\pi_{1+} - \pi_{+1} = [P(Drug = > ess) - P(placebo = > ess)]$$

There is 'marginal homogeneity' if $\pi_{1+} = \pi_{+1}$

Note:

$$\pi_{1+}-\pi_{+1}=(\pi_{11}+\pi_{12})-(\pi_{11}+\pi_{21})=\pi_{12}-\pi_{21}$$
 So, $\pi_{1+}=\pi_{+1} \iff \pi_{12}=\pi_{21}$ (Symmetry)

Under H_0 : marginal homogeneity,

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

Each of $n^*=n_{12}+n_{21}$ observations has probability 1/2 of contributing to n_{12} , 1/2 of contributing to n_{21} .

$$n_{12} \sim bin(n^*, 1/2), mean = \frac{n^*}{2}, std. dev. = \sqrt{n^*(\frac{1}{2})(\frac{1}{2})}$$

By normal approximation to binomial, for large n^* ,

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

or,
$$Z^2 = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} \sim approximate \ \chi_1^2$$

called McNemar's Test.

ex) Revisit (*)

$$Z = \frac{n_{12} - n_{21}}{\sqrt{n_{12} + n_{21}}} = \frac{49 - 10}{\sqrt{49 + 10}} = 5.1$$

$$\Rightarrow \left(\frac{49 - 10}{\sqrt{49 + 10}}\right)^2 = 25.7797 (McNemar's \ test)$$

$$p-vlaue < 0.0001 \ \text{for} \ H_0: \pi_{1+} = \pi_{+\,1} \ vs. \ H_a: \pi_{1+} \neq \pi_{+\,1}.$$

Extremely strong evidence that probability of success is higher for drug than placebo,

C.I. for $\pi_{1+} - \pi_{+1}$

Estimate $\pi_{1+}-\pi_{+1}$ by $p_{1+}-p_{+1}$, difference of sample proportions.

$$Var(p_{1+}-p_{+\,1}) = \, Var(p_{1+}) + \, Var(p_{+\,1}) - 2 \, Cov(p_{1+}, \ p_{+\,1})$$

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

ex) Revisit (*)

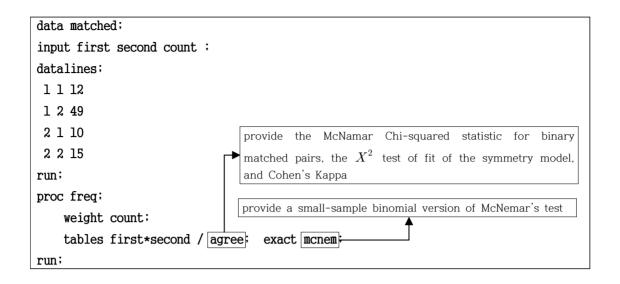
$$p_{1+} - p_{+1} = \frac{n_{11} + n_{12}}{n} - \frac{n_{11} + n_{21}}{n}$$
$$= \frac{n_{12} - n_{21}}{n} = \frac{49 - 10}{86} = 0.453$$

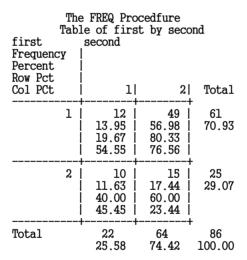
The standard error of $p_{1+}-p_{+\,1}$ is

$$\frac{1}{86}\sqrt{(49+10) - \frac{(49-10)^2}{86}} = 0.075$$

95% C.I. is $0.453 \pm 1.96(0.075) = (0.31, 0.60)$

Conclude we are 95% confident that probability of success is between 0.31 and 0.60 higher for drug than for placebo.





Statistics for Table of first by second

McNemar's T	est
Statistics(S) DF	25.7797 1
Asymptotic Pr > S	<.0001
Exact Pr >= S	<.0001

Simple Kappa Coefficient

Kappa		-0.1392
ASE		0.0792
95% Lower Conf	Limit	-0.2945
95% Upper Conf	Limit	0.0161

Sample Size = 86

Measuring agreement

ex) Movie reviews by Siskel and Ebert

		Ebert			
		Con.	Mixed	Pro.	
	Con.	24	8	13	45
Siskel	Mixed	8	13	11	45 32
	Pro.	10	9	64	83
		42	30	88	160

How strong is their agreement?

Let
$$\pi_{ij} = P(S=i, E=j)$$

$$P(agreement) = \pi_{11} + \pi_{22} + \pi_{33} = \sum_{i=1}^{3} \pi_{ii} = 1 \text{(if } perfect \ agreement)$$

If ratings are indep., $\pi_{ii} = \pi_{i+}\pi_{+i}$

Kappa

$$\begin{split} K &= \frac{\sum \! \pi_{ii} - \sum \! \pi_{i+} \pi_{+\,i}}{1 - \sum \! \pi_{i+} \pi_{+\,i}} \\ &= \frac{P(agree) - P(agree|independent)}{1 - P(agree|independent)} \end{split}$$

Note:

- \bullet K=0 if agreement only equals that expected under independence.
- K = 1 if perfect agreement
- Denominator = maximum difference for numerator, if perfect agreement.

ex)

$$\sum \hat{\pi}_{ii} = \frac{24 + 13 + 64}{160} = 0.63$$

$$\sum \hat{\pi}_{i+} \hat{\pi}_{+i} = \left(\frac{45}{160}\right) \left(\frac{42}{160}\right) + \dots + \left(\frac{83}{160}\right) \left(\frac{88}{160}\right) = 0.40$$

$$\hat{K} = \frac{0.63 - 0.40}{1 - 0.40} = 0.389$$

The strength of agreement is only moderate.

- 95% C.I. for $K : 0.389 \pm 1.96(0.06) = (0.27, 0.51)$
- lacktriangle For $H_0: K=0$,

$$Z = \frac{\hat{K}}{SE} = \frac{0.389}{0.06} = 6.5$$

There is extremely strong evidence that agreement is better than "Chance"

data movie; input siskel \$ ebert \$ count @@; cards; mixed 8 con con con 24 con pro 13 mixed con 8 mixed mixed 13 mixed pro 11 mixed 9 pro pro con 10 pro pro 64 run; proc freq data=movie; weight count; tables siskel*ebert / agree; exact mcnem; run;

The FREQ Procedure

테이블:siskel * ebert siskel ebert Frequency Percent Row Pct Col Pct mixed Total pro con 13 8.13 8 con 5.00 17.78 26.67 15.00 53.33 57.14 28.13 28.89 14.77 13 | 8.13 | 40.63 | 43.33 | 11 6.88 34.38 12.50 32 20.00 mixed 5.00 | 25.00 | 19.05 | pro 6.25 12.05 23.81 5.63 10.84 30.00 40.00 77.11 51.88 72.73 30 18.75 합계 42 26.25 160 55.00 100.00

Statistics for Table of siskel * ebert

Test of Symmetry -----Statistics(S) 0.5913 DF 3 Pr > S 0.8984

Statistic	Value	ppa Statist ASE	95% Confiden	ce Limits
Simple Kappa	0.3888	0.0598	0.2716	0.5060
Weighted Kappa	0.4269	0.0635	0.3024	0.5513

Sample Size = 160