# :Models for Matched Pairs:

The responses in the two samples are match pairs, if each observation in one sample pairs with an observation in the other sample. Because of the matching, the samples are statistically dependent, So methods that treat the two sets of observations as independent samples are inappropriate.

=> Examples of Matched Pairs:

★ A researcher attempts to determine if a drug has an effect on a particular disease. Counts of individuals are given in the following table, with the diagnosis (disease: present or absent) before treatment given in the rows and diagnosis after treatment in the columns. The same subjects to be included in the before-and-after measurements (matched pairs).

| | After | | |
|---|---|---|---|
| Before | Present | Absent | Row total. |
| Present | 101 | 121 | 222 |
| Absent | 59 | 33 | 92 |
| column total | 160 | 154 | 314 |

★ From the 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.

| Pay higher | cut living Standards | | |
|---|---|---|---|
| Taxes | Yes | No | row total |
| Yes | 227 | 132 | 359 |
| No | 107 | 678 | 785 |
| Column total | 334 | 810 | 1144 |

## Case-Control experiments:

| Case | Control: Exposed | Not exposed | row total |
|---|---|---|---|
| Exposed | 13 | 25 | 38 |
| Not exposed | 4 | 92 | 96 |
| Column total | 17 | 117 | 134 |

The investigator studies 134 cases and 134 matched controls (matched on the base of age, gender, occupation, location and other relevant variables), for a total of 268 subjects.

## Multinomial categorical response data for match pair

From the 2004 General Social Survey, reports respondents' region of residence in 2004 and at age 16.

| Residence at Age 16 | Residence in 2004: Northeast | Midwest | South | West |
|---|---|---|---|---|
| Northeast | 425 | 17 | 80 | 36 |
| Midwest | 10 | 155 | 74 | 47 |
| South | 7 | 34 | 771 | 33 |
| West | 5 | 14 | 29 | 452 |

## Ordinal categorical response data for matched pair.

From a General Social Survey. Subjects were asked "How often do you make a special effort to buy fruite and vegetables grown without pesticides or chemicals? and "How often do you make a special effort to sort glass or cans or plastic or papers and so on for recycling"?

| Chemical free | Recycle: Always | Sometimes | Never |
|---|---|---|---|
| Always | 66 | 39 | 3 |
| Sometimes | 227 | 359 | 48 |
| Never | 150 | 216 | 108 |

The most common way dependent samples occur is when each sample has the same subjects.

| Pay Higher taxes | Cut living Standards Yes | No | Total |
|---|---|---|---|
| Yes | 227 | 132 | 359 |
| No | 107 | 678 | 785 |
| Total | 334 | 810 | 1144 |

Comparing dependent proportions:
How can we compare the probabilies of a "Yes" outcome for the two environment questions?
Let $n_{ij}$ denote the number of subjects who respond in category i for the first question and j for the second question.

$n_{1+} = n_{11} + n_{12} = 359$ subjects said yes for raising taxes

$n_{+1} = n_{11} + n_{21} = 334$ subjects said yes for accepting cuts in living standards

The sample marginal proportion were:

$$P_{1+} = \frac{n_{1+}}{n} = \frac{359}{1144} = 0.31$$

$$P_{+1} = \frac{n_{+1}}{n} = \frac{334}{1144} = 0.29$$

These marginal proportions are correlated. We see that 227+ 678 subjects had the same opinion on both questions. The compose most of the sample since fewer people answered "Yes" on one and "no" on other. A strong association exists between the opinions on the two questions; the sample odd ratio being:

$$\text{Odd ration} = \frac{n_{11} \times n_{22}}{n_{12} \times n_{21}} = \frac{227 \times 678}{132 \times 107} = 10.9 > 1$$

There is association between the opinions on the two questions.

Scanned with CamScanner

⇒ Marginal homogeneity:
Let $\pi_{ij}$ denote the ~~probability of outcome i for 1 question~~ & j for 2 question
The Probability of a "Yes" outcome are $\pi_{1+}$ for question 1 and $\pi_{+1}$ for question 2. When these are identical, the probabilities of a "no" outcome are also identical. There is then said to be marginal homogeneity. Since

$$\boxed{\pi_{1+} = \pi_{+1}}$$

$$\pi_{1+} - \pi_{+1} = 0$$
$$(\pi_{11} + \pi_{12}) - (\pi_{11} + \pi_{21}) = 0$$
$$\cancel{\pi_{11}} + \pi_{12} - \cancel{\pi_{11}} - \pi_{21} = 0$$
$$\pi_{12} - \pi_{21} = 0$$
$$\boxed{\pi_{12} = \pi_{21}}$$

marginal homogeneity in 2x2 tables is equivalent to $\pi_{12} = \pi_{21}$.

⇒ **McNemar Test** {Test of marginal homogeneity / Comparing Marginal proportions:
The McNemar test is a non-parametric test for paired nominal data. It's used when you are interested in finding a change in proportions for the paired data. This test is sometimes reffered to as McNemar's Chi-square test because the test statistic has a chi-square distribution.
Assumptions for the McNemar test:
The three main assumptions for the test are:
1. You must have one nominal variable with two categories and one independent variable with two connected groups.
2- Participants cannot appear in more than one group
3- Your sample must be a random sample.

⇒ It is a test of marginal homogeneity for matched-pairs data with a binary response.

Considering the previous example (McNemar Test):

| Pay higher taxes | cut living Standards Yes | No | Total |
|---|---|---|---|
| Yes | 227 | 132 | 359 |
| No | 107 | 678 | 785 |
| Total | 334 | 810 | 1144 |

① Formulation of Hypothesis:

$H_0: \pi_{1+} = \pi_{+1}$ or $\pi_{12} = \pi_{21}$ (Marginal homogeneity)
$H_1: \pi_{1+} \neq \pi_{+1}$ or $\pi_{12} \neq \pi_{21}$ (Marginal hetrogeneity).

② Level of Significance.

$\alpha = 0.05$.

③ Test statistic:

$$\chi^2_{cal} = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} \quad \text{if } n_{12} + n_{21} > 25$$

This statistic has an approximate chi-squared distribution with $df = 1$

$$\left\{ \begin{aligned} df &= (r-1)(c-1) \\ &= (2-1)(2-1) \\ &= (1)(1) = 1. \end{aligned} \right\}$$

④ Calculation:

$$\chi^2_{cal} = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} = \frac{(132 - 107)^2}{132 + 107} = \frac{625}{239} = 2.62$$

⑤ Critical region:

$$\chi^2_{\alpha, 1} = 3.84$$

If $\chi^2_{cal} \geq 3.84$ then reject $H_0$ otherwise do not reject $H_0$.

⑥ Conclusion: There is significant change in the opinions of raising taxes and accepting cuts in living Standards.

Scanned with CamScanner