

# Categorical Data Analysis

## Chapter 10

Deyuan Li  
School of Management  
Fudan University

Fall 2015

## 10. Models for Matched Pairs

In this chapter, we introduce methods for comparing categorical responses for two samples when each observation in one sample pairs with an observation in the other.

For easy understanding, we assume  $n$  independent subjects and let

$$\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{it_i})$$

is the observation of subject  $i$  at different time.

## 10. Models for Matched Pairs

In statistics,  $\{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n\}$  are called **longitudinal data** (纵向数据); in econometrics it is called **panel data** (面板数据).

For fixed  $i$ ,  $\mathbf{Y}_i$  is a time series; for fixed time  $j$ ,  $\{Y_{1j}, Y_{2j}, \dots, Y_{nj}\}$  is a sequence of independent random variables.

If  $t_i = 2$  for all  $i$ ,  $\{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n\}$  is called **matched-pairs data** (配对数据). Note that the two samples  $\{Y_{11}, Y_{21}, \dots, Y_{n1}\}$  and  $\{Y_{12}, Y_{22}, \dots, Y_{n2}\}$  are **not independent**.

## 10. Models for Matched Pairs

**TABLE 10.1** Rating of Performance of Prime Minister

First Survey	Second Survey		Total
	Approve	Disapprove	
Approve	794	150	944
Disapprove	86	570	656
Total	880	720	1600

For a poll of a random sample of 1600 voting-age British citizens, 944 indicated approval of the Prime Minister's performance in office. Six months later, of these same 1600 people, 880 indicated approval.

For matched pairs with a categorical response, a two-way contingency table with the same row and column categories summarizes the data. The table is **square** (方阵) .

## 10. Models for Matched Pairs

Let  $n = 1600$  and  $\mathbf{Y}_i = (Y_{i1}, Y_{i2})$ , where

- $Y_{ij} = 1$  means approval of subject  $i$  in the  $j$ -th survey,
- $Y_{ij} = 0$  means disapproval of subject  $i$  in the  $j$ -th survey,

for  $i = 1, 2, \dots, n$  and  $j = 1, 2$ .

Then from Table 10.1, we have

$$\begin{aligned}\sum_i Y_{i1} &= 944, & \sum_i Y_{i2} &= 880, \\ \sum_i Y_{i1} Y_{i2} &= 794, & \sum_i (1 - Y_{i1})(1 - Y_{i2}) &= 570.\end{aligned}$$

A strong association exists between opinions six months apart, since the sample odds ratio is  $(794 \times 570)/(150 \times 86) = 35.1$ .

# Outline

- 1 Comparing Dependent Proportions
- 2 Conditional Logistic Regression for Binary Matched Pairs
- 3 Marginal Models for Squared Contingency Tables
- 4 Symmetry, Quasi-symmetry and Quasi-independence
- 5 Measurement Agreement between Observations
- 6 Bradley-Terry Model for Paired Preferences

# 10.1 Comparing Dependent Proportions

For each matched pair  $\mathbf{Y} = (Y_1, Y_2)$ , define

$$\pi_{ab} = P(Y_1 = a, Y_2 = b).$$

Assume  $\mathbf{Y}, \mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  are i.i.d. matched pairs.

Let  $n_{ab}$  count the number of such pairs, with  $p_{ab} = n_{ab}/n$  the sample proportion.

We treat  $\{n_{ab}\}$  as a sample from a multinomial  $(n; \{\pi_{ab}\})$  distribution.

Denote  $p_{a+}$  is the proportion in category  $a$  for observation 1, and  $p_{+a}$  is the corresponding proportion for observation 2.

# 10.1 Comparing Dependent Proportions

We compare samples by comparing marginal proportions  $\{p_{a+}\}$  with  $\{p_{+a}\}$ .

In this section we consider binary outcomes (i.e.  $a, b = 1, 2$ ). When  $\pi_{1+} = \pi_{+1}$ , then  $\pi_{2+} = \pi_{+2}$  also, and there is **marginal homogeneity**.

Since

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

marginal homogeneity in  $2 \times 2$  tables is equivalent to  $\pi_{12} = \pi_{21}$ . The table then shows *symmetry* across the main diagonal.



## 10.1.1 Inference for dependent proportions

Let  $d = p_{+1} - p_{1+} = p_{2+} - p_{+2}$ . Then (see Section 1.2.2)

$$\text{cov}(p_{+1}, p_{1+}) = (\pi_{11}\pi_{22} - \pi_{12}\pi_{21})/n.$$

Thus,

$$\text{var}(\sqrt{nd}) = \pi_{1+}(1 - \pi_{1+}) + \pi_{+1}(1 - \pi_{+1}) - 2(\pi_{11}\pi_{22} - \pi_{12}\pi_{21}).$$

For large samples,  $d$  has approximately a normal sampling distribution. A confidence interval for  $\delta = \pi_{+1} - \pi_{1+}$  is then

$$d \pm z_{\alpha/2} \hat{\sigma}(d),$$

where

$$\begin{aligned} \hat{\sigma}^2(d) &= [p_{1+}(1 - p_{1+}) + p_{+1}(1 - p_{+1}) - 2(p_{11}p_{22} - p_{12}p_{21})]/n \\ &= [(p_{12} + p_{21}) - (p_{12} - p_{21})^2]/n. \end{aligned}$$

## 10.1.1 Inference for dependent proportions

The hypothesis of marginal homogeneity is

$$H_0 : \pi_{1+} = \pi_{+1}, \quad (i.e. \delta = 0).$$

The Wald test statistic is  $z = d/\hat{\sigma}(d)$ ;

The score test statistic is  $z_0 = d/\hat{\sigma}_0(d)$ , where

$$\hat{\sigma}_0^2(d) = \frac{p_{12} + p_{21}}{n} = \frac{n_{12} + n_{21}}{n^2}.$$

$z_0$  can be simplified to

$$z_0 = \frac{n_{21} - n_{12}}{(n_{21} + n_{12})^{1/2}}$$

and  $z_0^2$  is an asymptotical chi-squared random with  $df=1$ .

The test using  $z_0$  is called *McNemar's test*.

## 10.1.2 Prime minister approval rating example

For Table 10.1,  $p_{1+} = 944/1600 = 0.59$  (approval in the first survey) and  $p_{+1} = 880/1600 = 0.55$  (approval in the second survey). A 95% CI for  $\pi_{+1} - \pi_{1+}$  is

$$(0.55 - 0.59) \pm 1.96(0.0095) = (-0.06, -0.02).$$

⇒ The approval rating appears to have dropped between 2% and 6%.

For testing marginal homogeneity, the score test statistic

$$z_0 = \frac{86 - 150}{(86 + 150)^{1/2}} = -4.17.$$

It shows strong evidence of a **drop** in the approval rating.

## 10.1.3 Increased precision with dependent samples

For two independent samples,

$$\begin{aligned}\text{var}[\sqrt{n}(\text{difference of sample proportions})] \\ = \pi_1(1 - \pi_1) + \pi_2(1 - \pi_2).\end{aligned}$$

For two dependent samples, one extra item  $-2(\pi_{11}\pi_{22} - \pi_{12}\pi_{21})$  is added.

If dependent samples exhibits a **positive dependence** with  $\log \theta = \log[\pi_{11}\pi_{22}/\pi_{12}\pi_{21}] > 0$ , then  $\text{var}(d)$  is **smaller than** when the samples are independent.

# Outline

- 1 Comparing Dependent Proportions
- 2 Conditional Logistic Regression for Binary Matched Pairs**
- 3 Marginal Models for Squared Contingency Tables
- 4 Symmetry, Quasi-symmetry and Quasi-independence
- 5 Measurement Agreement between Observations
- 6 Bradley-Terry Model for Paired Preferences

## 10.2.1 Marginal versus conditional models for matched pairs

Let  $(Y_1, Y_2)$  be a matched pair and  $Y_t$  be a binary random (1: *success*; 0: *failure*). Then the difference  $\delta = P(Y_2 = 1) - P(Y_1 = 1)$  occurs as a parameter in

$$P(Y_t = 1) = \alpha + \delta x_t, \quad (10.5)$$

where  $x_1 = 0$  and  $x_2 = 1$ . Alternatively, the logit link yields

$$\text{logit}[P(Y_t = 1)] = \alpha + \beta x_t. \quad (10.6)$$

The parameter  $\beta$  is a log odds ratio with the marginal distributions. The ML estimator of  $\beta$  is  $\hat{\beta} = \log[p_{+1}p_{2+}/p_{+2}p_{1+}]$  and see Problem 10.26 for its asymptotic variance.

Models (10.5) and (10.6) are called **marginal models** (边缘模型) since they focus on the marginal distributions of responses.

## 10.2.1 Marginal versus conditional models for matched pairs

Let  $(Y_{i1}, Y_{i2})$  denote the  $i$ th pair of observations. A model then has the form

$$\text{link}[P(Y_{it} = 1)] = \alpha_i + \beta \mathbf{x}_t. \quad (10.7)$$

This is called a **conditional model** (条件模型), since the effect  $\beta$  is defined conditional on the subject. This effect is **subject-specific**, since it is defined at the subject level.

With logit link,

$$P(Y_{it} = 1) = \frac{\exp(\alpha_i + \beta \mathbf{x}_t)}{1 + \exp(\alpha_i + \beta \mathbf{x}_t)}.$$

## 10.2.2 A logit model with subject-specific probabilities

Cox (1958, 1970) and Rasch(1961) presented model (10.7) with logit link:

$$\text{logit}[P(Y_{it} = 1)] = \alpha_i + \beta x_t. \quad (10.8)$$

$\Rightarrow$  common effect  $\beta$ . Obviously,

$$P(Y_{i1} = 1) = \frac{\exp(\alpha_i)}{1 + \exp(\alpha_i)}, \quad P(Y_{i2} = 1) = \frac{\exp(\alpha_i + \beta)}{1 + \exp(\alpha_i + \beta)}.$$

$\Rightarrow$  For each subject, the odds of success for observation 2 are  $\exp(\beta)$  times the odds for observation 1.



## 10.2.3 Conditional ML inference for binary matched pairs

For model (10.8), with assuming independence of responses for different subjects and for the two observations on the same subject (*note: we have to assume some association between two observations; here independence is assumed*), the joint mass function for  $\{(y_{11}, y_{12}), \dots, (y_{n1}, y_{n2})\}$  is

## 10.2.3 Conditional ML inference for binary matched pairs

$$\prod_{i=1}^n \left( \frac{\exp(\alpha_i)}{1 + \exp(\alpha_i)} \right)^{y_{i1}} \left( \frac{1}{1 + \exp(\alpha_i)} \right)^{1-y_{i1}} \\ \times \left( \frac{\exp(\alpha_i + \beta)}{1 + \exp(\alpha_i + \beta)} \right)^{y_{i2}} \left( \frac{1}{1 + \exp(\alpha_i + \beta)} \right)^{1-y_{i2}}.$$

In term of this data, it is proportional to

$$\exp \left[ \sum_i \alpha_i (y_{i1} + y_{i2}) + \beta \left( \sum_i y_{i2} \right) \right].$$

To eliminate  $\{\alpha_i\}$ , we condition on their **sufficient statistics**, the pairwise success totals  $\{S_i = y_{i1} + y_{i2}\}$ . Given  $S_i = 0$ ,  $P(Y_{i1} = Y_{i2} = 0) = 1$ , and given  $S_i = 2$ ,  $P(Y_{i1} = Y_{i2} = 1) = 1$ .

## 10.2.3 Conditional ML inference for binary matched pairs

The distribution of  $(Y_{i1}, Y_{i2})$  depends on  $\beta$  only when  $S_i = 1$ ; that is, only when outcomes differ for the two responses.

Given  $y_{i1} + y_{i2} = 1$ , the conditional distribution is

$$\begin{aligned}
 &P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | S_i = 1) \\
 &= P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}) / [P(Y_{i1} = 1, Y_{i2} = 0) + P(Y_{i1} = 0, Y_{i2} = 1)] \\
 &= \frac{\left(\frac{\exp(\alpha_i)}{1 + \exp(\alpha_i)}\right)^{y_{i1}} \left(\frac{1}{1 + \exp(\alpha_i)}\right)^{1-y_{i1}} \left(\frac{\exp(\alpha_i + \beta)}{1 + \exp(\alpha_i + \beta)}\right)^{y_{i2}} \left(\frac{1}{1 + \exp(\alpha_i + \beta)}\right)^{1-y_{i2}}}{\frac{\exp(\alpha_i)}{1 + \exp(\alpha_i)} \frac{1}{1 + \exp(\alpha_i + \beta)} + \frac{1}{1 + \exp(\alpha_i)} \frac{\exp(\alpha_i + \beta)}{1 + \exp(\alpha_i + \beta)}} \\
 &= \exp(\beta) / [1 + \exp(\beta)], \quad y_{i1} = 0, \quad y_{i2} = 1 \\
 &= 1 / [1 + \exp(\beta)], \quad y_{i1} = 1, \quad y_{i2} = 0.
 \end{aligned}$$

## 10.2.3 Conditional ML inference for binary matched pairs

Conditional on  $S_i = 1$ , the joint distribution of the matched pairs is

$$\prod_{S_i=1} \left( \frac{1}{1 + \exp(\beta)} \right)^{y_{i1}} \left( \frac{\exp(\beta)}{1 + \exp(\beta)} \right)^{y_{i2}} = \frac{[\exp(\beta)]^{n_{21}}}{[1 + \exp(\beta)]^{n_{12} + n_{21}}}.$$

Differentiating the log of this conditional likelihood, setting it to zero and solving the equation, we obtain the estimator of  $\beta$  and its standard error:

$$\hat{\beta} = \log(n_{21}/n_{12}), \quad \text{SE}(\hat{\beta}) = \sqrt{1/n_{12} + 1/n_{21}}.$$

## 10.2.6 Conditional ML for matched pairs with multiple predictors

When the binary response has  $p$  predictors for matched pairs, the model generalizes to

$$\text{logit}[P(Y_{it} = 1)] = \alpha_i + \beta_1 x_{1it} + \beta_2 x_{2it} + \cdots + \beta_p x_{pit},$$

where  $x_{hit}$  is the value of predictor  $h$  for observation  $t$  in pair.

Let  $\mathbf{x}_{it} = (x_{1it}, \dots, x_{pit})'$  and  $\beta = (\beta_1, \dots, \beta_p)'$ . A generalization of the derivation in Section 10.2.3 shows that

$$P(Y_{i1} = 0, Y_{i2} = 1 | S_i = 1) = \exp(\mathbf{x}'_{i2}\beta) / [\exp(\mathbf{x}'_{i1}\beta) + \exp(\mathbf{x}'_{i2}\beta)]$$

$$P(Y_{i1} = 1, Y_{i2} = 0 | S_i = 1) = \exp(\mathbf{x}'_{i1}\beta) / [\exp(\mathbf{x}'_{i1}\beta) + \exp(\mathbf{x}'_{i2}\beta)].$$

Similarly, we can estimate  $\beta$ .

# Outline

- 1 Comparing Dependent Proportions
- 2 Conditional Logistic Regression for Binary Matched Pairs
- 3 Marginal Models for Squared Contingency Tables**
- 4 Symmetry, Quasi-symmetry and Quasi-independence
- 5 Measurement Agreement between Observations
- 6 Bradley-Terry Model for Paired Preferences

## 10.3 Marginal Models for Squared Contingency Tables

We extend  $2 \times 2$  table to  $I \times I$  table with  $I > 2$ .

$n, n_{ab}, p_{ab}$  and  $(Y_1, Y_2)$  are the similar meaning as before. Let  $\pi_{ab} = P(Y_1 = a, Y_2 = b)$ .

Marginal homogeneity is

$$P(Y_1 = a) = P(Y_2 = a) \quad \text{for } a = 1, 2, \dots, I.$$

Marginal models compare  $\{P(Y_1 = a)\}$  and  $\{P(Y_2 = a)\}$ .

## 10.3.1 Marginal models for ordinal classification

For ordered categories, marginal model (10.6), for binary matched pairs extends using ordinal logits.

With cumulative logits,

$$\text{logit}[P(Y_t \leq j)] = \alpha_j + \beta x_t, \quad (10.14)$$

where  $t = 1, 2, j = 1, 2, \dots, I - 1, x_1 = 0$  and  $x_2 = 1$ .

This model has **proportional odds** structure.

$\Rightarrow$  The odds of outcome  $Y_2 \leq j$  equal  $\exp(\beta)$  times the odds of outcome  $Y_1 \leq j$ .



## 10.3.1 Marginal models for ordinal classification

The model implies **stochastically ordered** marginal distributions, with  $\beta > 0$  meaning that  $Y_1$  tends to be higher than  $Y_2$ .

Marginal homogeneity corresponds to  $\beta = 0$ .

Since the model refers to marginal probabilities  $\{P(Y_1 = a) = \pi_{a+}\}$  and  $\{P(Y_2 = b) = \pi_{+b}\}$ , one cannot substitute the model formula in the kernel  $\sum_a \sum_b n_{ab} \log \pi_{ab}$  of the log likelihood, which has joint probabilities.

⇒ We defer discussion of ML model fitting of marginal models to Section 11.2.5.

## 10.3.2 Premarital and extramarital sex example

**TABLE 10.5 Opinions on Premarital Sex and Extramarital Sex**

Premarital Sex	Extramarital Sex				Total
	1	2	3	4	
1	144	2	0	0	146
2	33	4	2	0	39
3	84	14	6	1	105
4	126	29	25	5	185
Total	387	49	33	6	475

*Source:* 1989 General Social Survey, National Opinion Research Center.

For a General Social Survey, subjects gave their opinion about premarital sex and extramarital sex. The response categories are 1=always wrong, 2=almost always wrong, 3=wrong only sometimes, 4=not wrong at all.

## 10.3.2 Premartial and extramarital sex example

The sample cumulative marginal proportions are (0.307, 0.389, 0.611) for premarital sex and (0.815, 0.918, 0.987) for extramarital sex.

⇒ This suggests that responses on premarital sex tended to be higher on the ordinal scale than those on extramarital sex.

With scores (1, 2, 3, 4), the mean for premarital sex is 2.69, closest to the "wrong only sometimes" score, and the mean response for extramarital sex is 1.28, closest to the "always wrong" score.

## 10.3.2 Premartial and extramartial sex example

The cumulative logit model (10.14) has  $\hat{\beta} = 2.51$  (SE=0.13). There is strong evidence that population responses are more positive on premarital than on extramarital sex.

The fit of the marginal homogeneity model has  $G^2 = 348.1$  (df=3), and the fit of model (10.14) has  $G^2 = 35.1$  (df=2). The ordinal model does not fit well, but it fits much better than the marginal homogeneity model.

Section 10.4.7 considers other better models.

## 10.3.3 Marginal models for nominal classifications

A baseline-category logit model has form

$$\log[P(Y_t = j)/P(Y_t = I)] = \alpha_j + \beta_j \mathbf{x}_t, \quad (10.15)$$

where  $t = 1, 2, j = 1, 2, \dots, I - 1$ ,  $x_1 = 0$  and  $x_2 = 1$ .

This model has  $2(I - 1)$  parameters for the  $2(I - 1)$  marginal probabilities. So it is saturated.

Marginal homogeneity is the special case  $\beta_1 = \dots = \beta_{I-1} = 0$ .

Three methods to test it.

## 10.3.4 Migration example

**TABLE 10.6 Migration from 1980 to 1985, with Fit of Marginal Homogeneity Model**

Residence in 1980	Residence in 1985				Total
	Northeast	Midwest	South	West	
Northeast	11,607 (11,607)	100 (98.1)	366 (265.7)	124 (94.0)	12,197 (12,064.7)
Midwest	87 (88.7)	13,677 (13,677)	515 (379.1)	302 (323.3)	14,581 (14,377.1)
South	172 (276.5)	255 (350.8)	17,819 (17,819)	270 (287.3)	18,486 (18,733.5)
West	63 (92.5)	176 (251.3)	286 (269.8)	10,192 (10,192)	10,717 (10,805.6)
Total	11,929 (12,064.7)	14,178 (14,377.1)	18,986 (18,733.5)	10,888 (10,805.6)	55,981

*Source:* Data based on Table 12 of U.S. Bureau of the Census, Current Population Reports, Series P-20, No. 420, *Geographical Mobility: 1985* (Washington, DC: U.S. Government Printing Office), 1987.

## 10.3.4 Migration example

Relatively few people changed region, 95% of the observations falling on the main diagonal. The ML fit of marginal homogeneity, shown in Table 10.6, gives  $G^2 = 240.8$  ( $df = 3$ ).

The sample marginal proportions for the four regions were  
 (0.218, 0.260, 0.330, 0.191) in 1980 and  
 (0.213, 0.253, 0.339, 0.194) in 1985.

Little change occurred.

To estimate the change for a given region, we apply (10.2) to the collapsed  $2 \times 2$  table that combines the other regions. A 95% CI for  $\pi_{+1} - \pi_{1+}$  is  $(0.2131 - 0.2179) \pm 1.96(0.00054)$ , or  $-0.005 \pm 0.001$ . Similarly, a 95% CI for  $\pi_{+2} - \pi_{2+}$  is  $-0.007 \pm 0.001$ , for  $\pi_{+3} - \pi_{3+}$  is  $-0.009 \pm 0.001$ , and  $\pi_{+4} - \pi_{4+}$  is  $-0.003 \pm 0.001$ . Although strong evidence of change occurs for all four regions, the changes were small.

# Outline

- 1 Comparing Dependent Proportions
- 2 Conditional Logistic Regression for Binary Matched Pairs
- 3 Marginal Models for Squared Contingency Tables
- 4 Symmetry, Quasi-symmetry and Quasi-independence**
- 5 Measurement Agreement between Observations
- 6 Bradley-Terry Model for Paired Preferences



## 10.4 Symmetry, Quasi-symmetry and Quasi-independence

An  $I \times I$  joint distribution  $\{\pi_{ab}\}$  satisfies **symmetry** (对称) if

$$\pi_{ab} = \pi_{ba} \quad \text{whenever } a \neq b. \quad (10.17)$$

Under symmetry,  $\pi_{a+} = \pi_{+a}$  for all  $a \Rightarrow$  **marginal homogeneity** (边缘齐次).

For  $I = 2$ , symmetry is equivalent to marginal homogeneity, but for  $I > 2$ , marginal homogeneity can occur without symmetry.

## 10.4.1 Symmetry as logit and loglinear models

When all  $\pi_{ab} > 0$ , symmetry is a logit and a loglinear model.

In logit form, it is trivially

$$\log(\pi_{ab}/\pi_{ba}) = 0, \quad \text{for } a < b.$$

For expected frequencies  $\{\mu_{ab} = n\pi_{ab}\}$ , it has the loglinear form

$$\log \mu_{ab} = \lambda + \lambda_a + \lambda_b + \lambda_{ab}, \quad (10.18)$$

where all  $\lambda_{ab} = \lambda_{ba}$ .

$\Rightarrow$  It can be simplified to  $\log \mu_{ab} = \lambda_{ab}$  with all  $\lambda_{ab} = \lambda_{ba}$ .

## 10.4.1 Symmetry as logit and loglinear models

For Poisson or multinomial cell counts  $\{n_{ab}\}$ , the likelihood equations are

$$\begin{aligned}\hat{\mu}_{ab} + \hat{\mu}_{ba} &= n_{ab} + n_{ba}, & \text{for all } a < b; \\ \hat{\mu}_{aa} &= n_{aa}, & \text{for all } a.\end{aligned}$$

The solution that satisfies symmetry is

$$\hat{\mu}_{ab} = \frac{n_{ab} + n_{ba}}{2} \quad \text{for all } a, b.$$

## 10.4.1 Symmetry as logit and loglinear models

The logit symmetry model has no parameters for the  $\binom{I}{2}$  binomial pairs  $\{(n_{ab}, n_{ba})\}$  with  $a < b$ , so its residual  $df = I(I - 1)/2$ . Equivalently, the loglinear symmetry model  $\log \mu_{ab} = \lambda_{ab}$  ( $\lambda_{ab} = \lambda_{ba}$ ) for  $I^2$  Poisson counts  $\{n_{ab}\}$  has  $\binom{I}{2}$   $\{\lambda_{ab}\}$  with  $a < b$  and  $I \setminus \{\lambda_{aa}\}$ , so  $df = I^2 - [I + I(I - 1)/2] = I(I - 1)/2$ .

For testing symmetry, Bowker (1948) showed that

$$\chi^2 = \sum_{a < b} \frac{(n_{ab} - n_{ba})^2}{n_{ab} + n_{ba}}.$$

For  $I = 2$  this is McNemar's statistic. The standardized Pearson residuals equal

$$r_{ab} = (n_{ab} - n_{ba}) / (n_{ab} + n_{ba})^{1/2}.$$

## 10.4.2 Quasi-symmetry

The loglinear model

$$\log \mu_{ab} = \lambda + \lambda_a^X + \lambda_b^Y + \lambda_{ab}, \quad (10.19)$$

where  $\lambda_{ab} = \lambda_{ba}$  for all  $a < b$ , is called **quasi-symmetry** (拟对称). It is different with (10.18).

$\Rightarrow$  symmetry if  $\lambda_a^X = \lambda_a^Y$  for  $a = 1, 2, \dots, I$ ;

$\Rightarrow$  independence if  $\lambda_{ab} = 0$  for all  $a$  and  $b$ .

The likelihood equations for quasi-symmetry are

$$\begin{aligned} \hat{\mu}_{a+} &= n_{a+}, & a &= 1, 2, \dots, I \\ \hat{\mu}_{+b} &= n_{+b}, & b &= 1, 2, \dots, I \\ \hat{\mu}_{ab} + \hat{\mu}_{ba} &= n_{ab} + n_{ba}, & \text{for } a &\leq b. \end{aligned} \quad (10.20)$$

Note: 1) only the first two sets of equations is needed;

2)  $\hat{\mu}_{aa} = n_{aa}$ ; 3) the residual  $df = (I - 1)(I - 2)/2$ .

## 10.4.2 Quasi-symmetry

The quasi-symmetry model has multiplicative form

$$\pi_{ab} = \alpha_a \beta_b \gamma_{ab}, \quad \text{where } \gamma_{ab} = \gamma_{ba} \quad \text{all } a < b \quad (10.21)$$

and all parameters are positive.

- It is a symmetry model when  $\alpha_a = \beta_a$  for all  $a$ .
- A table satisfying quasi-symmetry is the cellwise product of a table satisfying independence with one satisfying symmetry.
- The association symmetry implies that odds ratios on one side of the main diagonal are identical to corresponding odds ratios on the other side.

## 10.4.2 Quasi-symmetry

- In fact, the model can be defined by properties such as

$$\frac{\mu_{ab}\mu_{ll}}{\mu_{al}\mu_{lb}} = \frac{\mu_{ba}\mu_{ll}}{\mu_{bl}\mu_{la}} \quad \text{for all } a < b$$

or  $\theta_{ab} = \theta_{ba}$  for local odds ratios.

Goodman (1979a) referred to it as the symmetric association model.

## 10.4.3 Quasi-independence

A square contingency table satisfies **quasi-independence** (拟独立) when the variables are independent, given that the row and column outcomes differ. This has the loglinear form

$$\log \mu_{ab} = \lambda + \lambda_a^X + \lambda_b^Y + \delta_a I(a = b), \quad (10.23)$$

where  $I(\cdot)$  is the indicator function.

- This adds a parameter to the independence model for each cell on the main diagonal.
- The first three terms in (10.23) specify independence, and  $\delta_a$  permit  $\{\mu_{aa}\}$  to depart from this pattern and have arbitrary positive values.
- When  $\delta_a > 0$ ,  $\mu_{aa}$  is larger than under independence.



## 10.4.3 Quasi-independence

The likelihood equations for quasi-independence are

$$\hat{\mu}_{a+} = n_{a+}, \quad \hat{\mu}_{+a} = n_{+a}, \quad \hat{\mu}_{aa} = n_{aa}.$$

The model has  $I$  more parameters than the independence model, so its residual  $df = (I - 1)^2 - I$ . It applies to tables with  $I \geq 3$ .

Quasi-independence is the special case of quasi-symmetry (10.21) in which  $\{\gamma_{ab} \text{ for } a \neq b\}$  are identical.

Caussinus (1966) showed that they are equivalent when  $I = 3$ .

# 10.4.4 Migration revisited

**TABLE 10.7 Fit of Models to Table 10.6**

Residence in 1980	Residence in 1985 <sup>a</sup>				Total
	Northeast	Midwest	South	West	
Northeast	11,607	100 (126.6) <sup>1</sup> (95.8) <sup>2</sup>	366 (312.9) (370.4)	124 (150.5) (123.8)	12,197
Midwest	87 (117.4) (91.2)	13,677	515 (531.1) (501.7)	302 (255.5) (311.1)	14,581
South	172 (133.2) (167.6)	255 (243.8) (238.3)	17,189	270 (290.0) (261.1)	18,486
West	63 (71.4) (63.2)	176 (130.6) (166.9)	286 (323.0) (294.9)	10,192	10,717
Total	11,929	14,178	18,986	10,888	55,981

<sup>a1</sup>Quasi-independence fit; <sup>2</sup>quasi-symmetry fit; both models giving perfect fit on main diagonal.

## 10.4.4 Migration revisited

Not surprisingly,

- the independence model fits terribly, with  $G^2 = 125,923$  and  $X^2 = 146,929$ . (The maximum possible value of  $X^2$  is  $3n = 167,943$ ; see Problem 3.33).
- The symmetry model is also unpromising.  
For instance, 124 people moved from the northeast to the west, but only 63 people made the reverse move. The deviance for testing symmetry is  $G^2 = 243.6$  ( $df = 6$ ).

## 10.4.4 Migration revisited

Quasi-independence states that for people who moved, residence in 1985 is independent of region in 1980.

- $G^2 = 69.5$  ( $df = 5$ ).  $\Rightarrow$  This model fits much better than the independence model, primarily because it forces a perfect fit on the main diagonal, where most observations occur.
- Lack of fit is apparent off that diagonal. More people moved from the northeast to the south and fewer moved from the west to the south than quasi-independence predicts.

## 10.4.4 Migration revisited

The quasi-symmetry model has  $G^2 = 3.0$ , with  $df = 3$ .

⇒ It is much better than with quasi-independence.

The lack of symmetry in cell probabilities reflects slight marginal heterogeneity.

The subject-specific effects can be described using the model's parameter estimates,

$$\{\hat{\lambda}_1^Y - \hat{\lambda}_1^X = -0.672, \hat{\lambda}_2^Y - \hat{\lambda}_2^X = -0.623, \hat{\lambda}_3^Y - \hat{\lambda}_3^X = 0.122\}.$$

For instance, for a given subject the estimated odds of living in the south instead of the west in 1985 were  $\exp(0.122) = 1.13$  times the odds in 1980.

**Summary.** Quasi-symmetry model is the best.

## 10.4.5 Marginal homogeneity and quasi-symmetry

Marginal homogeneity is not equivalent to a loglinear model. However, quasi-symmetry is a useful model for studying marginal homogeneity.

Two results:

- Caussinus (1966) showed that symmetry is equivalent to quasi-symmetry and marginal homogeneity holding simultaneously.
- We also show (see p. 428) the converse, that the joint occurrence of quasi-symmetry and marginal homogeneity implies symmetry.

Notation: quasi-symmetry (QS); symmetry (S);  
marginal homogeneity (MH).

## 10.4.5 Marginal homogeneity and quasi-symmetry

When the quasi-symmetry (QS) holds, marginal homogeneity (MH) is equivalent to symmetry (S).

⇒ Conditional on quasi-symmetry, testing marginal homogeneity is equivalent to testing symmetry.

A test of marginal homogeneity compares fit statistics for the symmetry and quasi-symmetry models,

$$G^2(S|QS) = G^2(S) - G^2(QS),$$

with  $df = I - 1$ . Alternative approach using marginal models.

## 10.4.5 Marginal homogeneity and quasi-symmetry

Table 10.6 on migration from 1980 to 1985 has  $G^2(S) = 243.6$  and  $G^2(QS) = 3.0$ . The difference  $G^2(Q|QS) = 240.6$  ( $df = 3$ ). shows extremely strong evidence of marginal heterogeneity.

⇒ Results are similar to those quoted in Section 10.3.4 for the likelihood-ratio test based on model (10.15), for which  $G^2 = 240.8$ , or the Wald test, for which  $W = 236.5$  (both with  $df = 3$ ).



## 10.4.6 Ordinal quasi-symmetry model

For ordered categories, let  $u_1 \leq u_2 \leq \dots \leq u_I$ . An ordinal quasi-symmetry model is

$$\log \mu_{ab} = \lambda + \lambda_a + \lambda_b + \beta u_b + \lambda_{ab}, \quad (10.26)$$

where  $\lambda_{ab} = \lambda_{ba}$  for all  $a < b$ .

- It is a special case of quasi-symmetry model (10.19) in which  $\lambda_b^Y - \lambda_b^X = \beta u_b$  has a linear trend.
- Symmetry is the special case  $\beta = 0$ .
- This model has logit representation

$$\log(\pi_{ab}/\pi_{ba}) = \beta(u_b - u_a) \quad \text{for } a \leq b. \quad (10.27)$$

(10.27) is a special case of linear logit model.

## 10.4.6 Ordinal quasi-symmetry model

The likelihood equations for ordinal quasi-symmetry are

$$\sum_a u_a \hat{\mu}_{a+} = \sum_a u_a n_{a+}, \quad \sum_b u_b \hat{\mu}_{+b} = \sum_b u_b n_{+b},$$

$$\hat{\mu}_{ab} + \hat{\mu}_{ba} = n_{ab} + n_{ba} \quad \text{for } a < b.$$

When  $\beta = 0$ , this model implies stochastically ordered margins.

When  $\beta > 0$  ( $\beta < 0$ ), responses have a higher mean in the column (row) distribution.

## 10.4.6 Ordinal quasi-symmetry model

A test of marginal homogeneity ( $H_0 : \beta = 0$ ) uses  
ordinal quasi-symmetry + marginal homogeneity = symmetry.

The likelihood-ratio test statistic compares the deviance for symmetry and ordinal quasi-symmetry.

- fit this model by fitting (10.27) with logit model software:  
Identify  $(\pi_{ab}, \pi_{ba})$  as binomial with  $n_{ab} + n_{ba}$  trials, and fit a logit model with no intercept and predictor  $x = u_b - u_a$ .
- or fit (10.26) using iterative methods for loglinear models.

## 10.4.7 Premarital and extramarital sex revisited

For Table 10.5,

- the symmetry model is inadequate ( $G^2 = 402.2$ ,  $df = 6$ );
- quasi-symmetry fits well ( $G^2 = 1.4$ ,  $df = 3$ );
- the simpler model of ordinal quasi-symmetry also fits well: with scores  $\{1, 2, 3, 4\}$ ,  $G^2 = 2.1$  ( $df = 5$ ).

The ML estimate  $\hat{\beta} = 2.86$ . From (10.27), the estimated probability that outcome on premarital sex is  $x$  categories more positive than the outcome on extramarital sex equals  $\exp(0.86x)$  times the reverse probability.

# Outline

- 1 Comparing Dependent Proportions
- 2 Conditional Logistic Regression for Binary Matched Pairs
- 3 Marginal Models for Squared Contingency Tables
- 4 Symmetry, Quasi-symmetry and Quasi-independence
- 5 Measurement Agreement between Observations**
- 6 Bradley-Terry Model for Paired Preferences

## 10.5 Measurement Agreement between Observations

We now discuss an application, analyzing agreement between two observers, that uses matched-pairs models.

We illustrate with Table 10.8. This shows ratings by two pathologists (病理学家), labeled *A* and *B*, who separately classified 118 slides regarding the presence and extent of carcinoma of the uterine cervix (子宫颈癌).

The rating scale has the ordered categories

- 1 negative (阴性),
- 2 atypical squamous hyperplasia (异常鳞状增生),
- 3 carcinoma in situ (原位癌),
- 4 squamous or invasive carcinoma (鳞状或浸润性癌).

# 10.5 Measurement Agreement between Observations

**TABLE 10.8 Diagnoses of Carcinoma**

Pathologist A	Pathologist B <sup>a</sup>				Total
	1	2	3	4	
1	22 (8.5)	2 (-0.5)	2 (-5.9)	0 (-1.8)	26
2	5 (-0.5)	7 (3.2)	14 (-0.5)	0 (-1.8)	26
3	0 (-4.1)	2 (-1.2)	36 (5.5)	0 (-2.3)	38
4	0 (-3.3)	1 (-1.3)	17 (0.3)	10 (5.9)	28
Total	27	12	69	10	118

<sup>a</sup>Values in parentheses are standardized Pearson residuals for the independence model.

Source: N. S. Holmquist, C. A. McMahon, and O. D. Williams, *Arch. Pathol.* **84**: 334–345 (1967); reprinted with permission from the American Medical Association. See also Landis and Koch (1977).

## 10.5.1 Agreement: dependence and independence

Let  $\pi_{ab}$  denote the probability that observer  $A$  classifies a slide in category  $a$  and observer  $B$  classifies it in category  $b$ .

- $\pi_{aa}$  is the probability of both choosing category  $a$ ;
- $\sum_a \pi_{aa}$  is the total probability of agreement.
- Perfect agreement occurs when  $\sum_a \pi_{aa} = 1$ .

Our analysis focus on describing strength of agreement and detecting patterns of disagreement.



## 10.5.1 Agreement: dependence and independence

Agreement (一致) and association (关联) are distinct facets of the joint distribution.

- Strong agreement requires strong association,
- but strong association can exist without strong agreement.  
If observer  $A$  consistently rates subjects one category higher than observer  $B$ , strength of agreement is poor even though the association is strong.

## 10.5.1 Agreement: dependence and independence

**Evaluations of agreement:** compare  $\{n_{ab}\}$  to the values  $\{n_{a+}n_{+b}/n\}$  predicted under independence.

- That model is a baseline, showing the agreement expected if no association existed between ratings.
- Normally, it fits poorly if even mild agreement exists, but its cell standardized residuals show patterns of agreement and disagreement.
- Ideally, standardized residuals are large positive on the main diagonal and large negative off that diagonal.

## 10.5.1 Agreement: dependence and independence

The independence model fits Table 10.8 poorly ( $G^2 = 118.0$ ,  $df = 9$ ).

That table reports the standardized Pearson residuals in parentheses.

- The large positive residuals on the main diagonal indicate that agreement for each category is greater than expected by chance, especially for the first category.
- Off the main diagonal they are primarily negative.
- Disagreements occurred less than expected under independence, although the evidence of this is weaker for categories closer together.
- The most common disagreements were observer B choosing category 3 and observer A instead choosing category 2 or 4.

## 10.5.2 Using quasi-independence

Quasi-independence adds main-diagonal parameters  $\{\delta_a\}$ . For Table 10.8, this model has  $G^2 = 13.2$  ( $df = 5$ ). It fits much better than independence, but some lack of fit remains.

**TABLE 10.9 Fitted Values for Carcinoma Diagnoses of Table 10.8**

Pathologist $A$	Pathologist $B^a$			
	1	2	3	4
1	22	2	2	0
	(22) <sup>1</sup>	(0.7)	(3.3)	(0.0)
	(22) <sup>2</sup>	(2.4)	(1.6)	(0.0)
2	5	7	14	0
	(2.4)	(7)	(16.6)	(0.0)
	(4.6)	(7)	(14.4)	(0.0)
3	0		36	0
	(0.8)	(1.2)	(36)	(0.0)
	(0.4)	(1.6)	(36)	(0.0)
4	0	1	17	10
	(1.9)	(3.0)	(13.1)	(10)
	(0.0)	(1.0)	(17.0)	(10)

<sup>a1</sup>Quasi-independence model; <sup>2</sup>quasi-symmetry model.

## 10.5.2 Using quasi-independence

For two subjects, suppose that each observer classifies one in category  $a$  and one in category  $b$ . The odds that the observers agree rather than disagree on which is in category  $a$  and which is in category  $b$  equal

$$\tau_{ab} = \frac{\pi_{aa}\pi_{bb}}{\pi_{ab}\pi_{ba}} = \frac{\mu_{aa}\mu_{bb}}{\mu_{ab}\mu_{ba}}.$$

As  $\tau_{ab}$  increases, the observers are more likely to agree for that pair of categories. Under quasi-independence,

$$\tau_{ab} = \exp(\delta_a + \delta_b).$$

Larger  $\{\delta_a\}$  represent stronger agreement. For instance, for Table 10.8,  $\hat{\delta}_2 = 0.6$  and  $\hat{\delta}_3 = 1.9$ , and  $\hat{\tau}_{23} = 12.3$ . The degree of agreement seems fairly strong for other pairs of categories.

## 10.5.4 Kappa measure of agreement

A single index summarizes agreement. For **nominal** scales, the most popular measure is **Cohen's kappa**. It compares the probability of agreement  $\sum_a \pi_{aa}$  to that expected if the ratings were independent,  $\sum_a \pi_{a+} \pi_{+a}$ , by

$$\kappa = \frac{\sum_a \pi_{aa} - \sum_a \pi_{a+} \pi_{+a}}{1 - \sum_a \pi_{a+} \pi_{+a}}.$$

- $\kappa = 0$  when the agreement merely equals that expected under independence.
- $\kappa = 1.0$  when perfect agreement occurs.
- The stronger the agreement, the higher is  $\kappa$ , for given marginal distributions.
- $\kappa < 0$  occurs when agreement is weaker than expected by chance, but this rarely happens.

## 10.5.4 Kappa measure of agreement

For multinomial sampling, the sample value  $\hat{\kappa}$ :

$$\hat{\kappa} = \frac{\sum_a n_{aa}/n - \sum_a n_{a+}n_{+a}/n^2}{1 - \sum_a \sum_a n_{a+}n_{+a}/n^2}$$

has a large-sample normal distribution.

Its estimated asymptotic variance is

$$\hat{\sigma}^2(\hat{\kappa}) = \frac{1}{n} \left\{ \frac{P_o(1 - P_o)}{(1 - P_e)^2} + \frac{2(1 - P_o)[2P_oP_e - \sum_a p_{aa}(p_{a+} + p_{+a})]}{(1 - P_e)^3} \right. \\ \left. + \frac{(1 - P_o)^2[\sum_a \sum_b p_{ab}(p_{b+} + p_{+a})^2 - 4P_e^2]}{(1 - P_e)^4} \right\}$$

where  $P_o = \sum_a p_{aa}$  and  $P_e = \sum_a p_{a+}p_{+a}$ .

## 10.5.4 Kappa measure of agreement

It is more relevant to estimate strength of agreement by interval estimation of  $\kappa$ .

For Table 10.8,  $P_o = 0.636$  and  $P_e = 0.281$ . Sample kappa equals  $(0.636 - 0.281)/(1 - 0.281) = 0.493$ .

The difference between observed agreement and that expected under independence is about 50% of the maximum possible difference.

The estimated standard error is 0.057, so  $\kappa$  apparently falls roughly between 0.4 and 0.6, moderately strong agreement.



## 10.5.5 weighted kappa: quantifying disagreement

For **ordinal** scales, the popular measure is *weighted kappa* defined by

$$\kappa_w = \frac{\sum_a \sum_b w_{ab} \pi_{ab} - \sum_a \sum_b w_{ab} \pi_{a+} \pi_{+b}}{1 - \sum_a \sum_b w_{ab} \pi_{a+} \pi_{+b}},$$

where the weights  $\{w_{ab}\}$  satisfying  $0 \leq w_{ab} \leq 1$  with all  $w_{aa} = 1$  and all  $w_{ab} = w_{ba}$ .

- The weighted agreement is defined  $\sum_a \sum_b w_{ab} \pi_{ab}$ .
- One choice is  $w_{ab} = |a - b| / (I - 1)$ : agreement is greater for cells nearer the main diagonal.
- Fleiss and Cohen (1973) suggested  $w_{ab} = 1 - (a - b)^2 / (I - 1)^2$ .

# Outline

- 1 Comparing Dependent Proportions
- 2 Conditional Logistic Regression for Binary Matched Pairs
- 3 Marginal Models for Squared Contingency Tables
- 4 Symmetry, Quasi-symmetry and Quasi-independence
- 5 Measurement Agreement between Observations
- 6 Bradley-Terry Model for Paired Preferences**

## 10.6 Bradley-Terry Model for Paired Preferences

Sometimes, categorical outcomes result from **pairwise evaluations**.

A common example is athletic competitions, when the outcome for a team or player consists of categories (win, lose).

Another example is pairwise comparison of product brands, such as two brands of wine of some type.

When a wine critic rates  $I$  brands, it might be difficult to establish an outright ranking, especially if  $I$  is large. However, for any given pair, the critic could probably state a preference after tasting them at the same occasion.

An overall ranking of the wines could then be based on the pairwise preferences.

We present a model for this in this section.

## 10.6.1 Bradley-Terry model

Let  $\Pi_{ab}$  denote the probability that  $a$  is preferred to  $b$ .

Suppose that  $\Pi_{ab} + \Pi_{ba} = 1$  for all pairs; that is, a tie cannot occur.

The Bradley-Terry model is

$$\log \frac{\Pi_{ab}}{\Pi_{ba}} = \beta_a - \beta_b. \quad (10.31)$$

Alternatively,

$$\Pi_{ab} = \exp(\beta_a) / [\exp(\beta_a) + \exp(\beta_b)].$$

Thus,  $\Pi_{ab} = 1/2$  when  $\beta_a = \beta_b$  and  $\Pi_{ab} > 1/2$  when  $\beta_a > \beta_b$ .

## 10.6.1 Bradley-Terry model

Identifiability requires a constraint such as  $\beta_I = 0$  or  $\sum_a \exp(\beta_a) = 1$ .

Since the model describes  $\binom{I}{2}$  probabilities ( $\{\Pi_{ab}\}$  for  $a < b$ ) by  $(I - 1)$  parameters, residual  $df = \binom{I}{2} - (I - 1)$ .

For  $a < b$ , let

$N_{ab}$ : the sample number of evaluations on  $a$  and  $b$ ;

$n_{ab}$ : the sample number of evaluations:  $a$  is preferred to  $b$ ;

Construct a contingency table with empty cells on the main diagonal.

Assumption: evaluations for different pairs are independent.

## 10.6.2 Home team advantage in baseball

Table 10.10 shows results of the 1987 season for the seven baseball teams in the Eastern Division of the American League. For instance, of games between Boston and New York, Boston won 7 and New York won 6.

**TABLE 10.10 Results of 1987 Season for American League Baseball Teams**

Winning Team	Losing Team <sup>a</sup>						
	Milwaukee	Detroit	Toronto	New York	Boston	Cleveland	Baltimore
Milwaukee	—	7 (7.0)	9 (7.4)	7 (7.6)	7 (8.0)	9 (9.2)	11 (10.8)
Detroit	6 (6.0)	—	7 (7.0)	5 (7.1)	11 (7.6)	9 (8.8)	9 (10.5)
Toronto	4 (5.6)	6 (6.0)	—	7 (6.7)	7 (7.1)	8 (8.4)	12 (10.2)
New York	6 (5.4)	8 (5.9)	6 (6.3)	—	6 (7.0)	7 (8.3)	10 (10.1)
Boston	6 (5.0)	2 (5.4)	6 (5.9)	7 (6.0)	—	7 (7.9)	12 (9.8)
Cleveland	4 (3.8)	4 (4.2)	5 (4.6)	6 (4.7)	6 (5.1)	—	6 (8.6)
Baltimore	2 (2.2)	4 (2.5)	1 (2.8)	3 (2.9)	1 (3.2)	7 (4.4)	—

<sup>a</sup>Values in parentheses represent the fit of the Bradley-Terry model.

Source: *American League Red Book*, 1988 (St. Louis, MO: Sporting News Publishing Co.)

## 10.6.2 Home team advantage in baseball

We fitted the Bradley-Terry model as a logit model for  $\binom{I}{2}$  independent binomial samples, using an appropriate model matrix and no intercept (e.g., for SAS, see Table A.19).

The model fits adequately ( $G^2 = 15.6$ ,  $df = 15$ ). Table 10.10 contains the fitted values  $\hat{\mu}_{ab}$ .

(setting  $\hat{\beta}_7 = 0$ ) and setting  $\sum_a \exp(\hat{\beta}_a) = 1$  for Table 10.11)

**TABLE 10.11 Results of Fitting Bradley-Terry Models to Baseball Data**

Team	Winning Percentage	$\hat{\beta}_i$ (10.31)	$\exp(\hat{\beta}_i)$ (10.31)	$\exp(\hat{\beta}_i)$ (10.32)
Milwaukee	64.1	1.58	0.218	0.220
Detroit	60.2	1.44	0.189	0.190
Toronto	56.4	1.29	0.164	0.164
New York	55.1	1.25	0.158	0.157
Boston	51.3	1.11	0.136	0.137
Cleveland	39.7	0.68	0.089	0.088
Baltimore	23.1	0.00	0.045	0.044

## 10.6.2 Home team advantage in baseball

When Boston played New York, the estimated probability that Boston won is

$$\hat{\Pi}_{54} = \exp(\hat{\beta}_5) / [\exp(\hat{\beta}_5) + \exp(\hat{\beta}_4)] = 0.46.$$

The standard error of each  $\hat{\beta}_a$  and of each  $\hat{\beta}_a - \hat{\beta}_b$  is about 0.3, so not much evidence exists of a difference among the top five teams.

This model does not recognize which team is the home team. Most sports have a home field advantage: A team is more likely to win when it plays at its home city.

Table 10.12 contains results for the 1987 season according to the (home team, away team) classification.



## 10.6.2 Home team advantage in baseball

**TABLE 10.12 Wins / Losses by Home and Away Team, 1987**

Home Team	Away Team						
	Milwaukee	Detroit	Toronto	New York	Boston	Cleveland	Baltimore
Milwaukee	—	4-3	4-2	4-3	6-1	4-2	6-0
Detroit	3-3	—	4-2	4-3	6-0	6-1	4-3
Toronto	2-5	4-3	—	2-4	4-3	4-2	6-0
New York	3-3	5-1	2-5	—	4-3	4-2	6-1
Boston	5-1	2-5	3-3	4-2	—	5-2	6-0
Cleveland	2-5	3-3	3-4	4-3	4-2	—	2-4
Baltimore	2-5	1-5	1-6	2-4	1-6	3-4	—

*Source: American League Red Book, 1988 (St. Louis, MO: Sporting News Publishing Co.).*

For instance, when Boston was the home team, it beat New York 4 times and lost 2 times; when New York was the home team, it beat Boston 4 times and lost 3 times.

## 10.6.2 Home team advantage in baseball

Now for all  $a \neq b$ , let  $\Pi_{ab}^*$  denote the probability that team  $a$  beats team  $b$ , when  $a$  is the home team.

Consider logit model

$$\log \frac{\Pi_{ab}^*}{1 - \Pi_{ab}^*} = \alpha + (\beta_a - \beta_b). \quad (10.32)$$

When  $\alpha > 0$ , a home field advantage exists.

The home team of two evenly matched teams has probability  $\exp(\alpha)/[1 + \exp(\alpha)]$  of winning.

## 10.6.2 Home team advantage in baseball

For Table 10.12, model (10.32) describes 42 binomial distributions with 7 parameters. It has  $G^2 = 38.6$  ( $df = 35$ ).

Table 10.11 displays  $\{\exp(\hat{\beta}_a)\}$ , similar to those obtained previously. The estimate of the home-field parameter is  $\hat{\alpha} = 0.302$ .

For two evenly matched teams, the home team had estimated probability 0.575 of winning. When Boston played New York, the estimated probability of a Boston win was 0.54 at Boston and 0.39 at New York.

Model (10.32) is a useful generalization of the Bradley-Terry model whenever an order effect exists. For instance, in pairwise taste evaluations, the product tasted first may have a slight advantage.

# Appendix: SAS Codes

**TABLE A.16 SAS Code for McNemar's Test and Comparing Proportions for Matched Samples in Table 10.1**

---

```
data matched;
input first second count @@;
datalines;
1 1 794    1 2 150    2 1 86    2 2 570
;
proc freq; weight count;
  tables first*second/agree; exact mcnem;
proc catmod; weight count;
  response marginals;
  model first*second= (1 0 ,
                      1 1 ;
```

---

**TABLE A.17 SAS Code for Testing Marginal Homogeneity with Migration Data in Table 10.6**


---

```

data migrate;
input then $ now $ count m11 m12 m13 m21 m22 m23 m31 m32 m33 m44 m1 m2 m3;
datalines;
  ne ne 11607 1 0 0 0 0 0 0 0 0 0 0 0 0
  ne mw 100 0 1 0 0 0 0 0 0 0 0 0 0 0
  ne s 366 0 0 1 0 0 0 0 0 0 0 0 0 0
  ne w 124 -1 -1 -1 0 0 0 0 0 0 0 1 0 0
  mw ne 87 0 0 0 1 0 0 0 0 0 0 0 0 0
  mw mw 13677 0 0 0 0 1 0 0 0 0 0 0 0 0
  mw s 515 0 0 0 0 0 1 0 0 0 0 0 0 0
  mw w 302 0 0 0 -1 -1 -1 0 0 0 0 0 1 0
  s ne 172 0 0 0 0 0 0 1 0 0 0 0 0 0
  s mw 225 0 0 0 0 0 0 0 1 0 0 0 0 0
  s s 17819 0 0 0 0 0 0 0 0 1 0 0 0 0
  s w 270 0 0 0 0 0 0 -1 -1 -1 0 0 0 1
  w ne 63 -1 0 0 -1 0 0 -1 0 0 0 1 0 0
  w mw 176 0 -1 0 0 -1 0 0 -1 0 0 0 1 0
  w s 286 0 0 -1 0 0 -1 0 0 -1 0 0 0 1
  w w 10192 0 0 0 0 0 0 0 0 0 1 0 0 0
;
proc genmod;
  model count=m11 m12 m13 m21 m22 m23 m31 m32 m33 m44 m1 m2 m3
    / dist=poi link=identity;
proc catmod; weight count; response marginals;
  model then*now=_response_ /freq;
  repeated time 2;

```

---

**TABLE A.18 SAS Code Showing Square-Table Analysis of Table 10.5**


---

```

data sex;
input premar extramar symm qi count @@;
unif=premar*extramar;
datalines;
1 1 1 1 144      1 2 2 5   2      1 3 3 5   0      1 4   4 5   0
2 1 2 5   33      2 2 5 2   4      2 3 6 5   2      2 4   7 5   0
3 1 3 5   84      3 2 6 5 14      3 3 8 3   6      3 4   9 5   1
4 1 4 5 126      4 2 7 5 29      4 3 9 5 25      4 4 10 4 5
;
proc genmod; class symm;
  model count=symm/dist=poi link=log; * symmetry;
proc genmod; class extramar premar symm;
  model count=symm extramar premar/dist=poi link=log; *QS;
proc genmod; class symm;
  model count=symm extramar premar/dist=poi link=log; * ordinal QS;
proc genmod; class extramar premar qi;
  model count=extramar premar qi/dist=poi link=log; * quasi indep;
proc genmod; class extramar premar;
  model count=extramar premar unif/dist=poi link=log;
data sex2;
input score below above @@; trials=below+above;
datalines;
1 33 2      1 14 2      1 25 1      2 84 0      2 29 0      3 126 0
;
proc genmod data=sex2;
  model above/trials=score/dist=bin link=logit noint;
proc genmod data=sex2;
  model above/trials=/dist=bin link=logit noint;
proc genmod data=sex2;
  model above/trials=/dist=bin link=logit;

```

---