

8 Models for Matched Pairs

- This chapter introduces methods for comparing categorical responses for two samples when each sample has the same subject or when a natural pairing exists between each subject in one sample and a subject from the other sample.
- The responses from the two samples are *statistically dependent*. Methods that treat the two sets of observations as independent samples are not appropriate.

- As we discussed in Chapter 5, we need a large sample size relative to the number of parameters for the maximum likelihood estimates to be valid.

If the data are highly stratified (e.g., $2 \times 2 \times K$ and K is large) and there are few subjects in each stratum, the sample size is often inadequate for the usual logistic regression since you will need to estimate a large number of parameters for the stratification effects.

- Highly stratified data often come from a design with cluster sampling. These are designs where two or more observations are made on each primary sampling unit or cluster.

Example: Common examples of such data are paired observations such as fraternal twins, litter mates, before-and-after outcomes from the same subject, or two occasions for expressing an opinion.

- Matched case-control studies in epidemiology also produce highly stratified data. In these studies, you match cases (those subjects with a disease or condition) with controls (those subjects without the disease or condition) on the basis of possibly confounding variables such as age, race, and sex.
- Models for data from matched case-control studies contain a separate parameter for each matched pair, so the number of parameters increases with sample size.
- The parameters associated with the stratification effect are thought of as *nuisance parameters*.
- The appropriate form of logistic regression for these types of data with nuisance parameters is called *conditional logistic regression*. By conditioning arguments, the pair or matching effect is eliminated when carrying out statistical inference on the effects of interest.
- In section 8.1, we compare proportions from two dependent samples. In section 8.2, we carry out logistic regression for matched pairs with binary responses.

8.1 Comparing Dependent Proportions

We discuss methods of comparing dependent proportions for binary response variables.

Example: **Rating of Performance of the Prime Minister**

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

- n_{ij} = the number of subjects making response i to the first survey and response j to the second survey
- $n_{1+} = n_{11} + n_{12} = 944$ is the number of approve on the first survey.
- $n_{+1} = n_{11} + n_{21} = 880$ is the number of approve on the second survey.
- The sample proportions of $p_{1+} = 944/1600 = 0.59$ and $p_{+1} = 880/1600 = 0.55$ are correlated, so the data analysis must take this into account.

- Let π_{ij} = the probability that a subject makes response i at survey 1 and response j at survey 2.
- Then the probabilities of approval are π_{1+} for survey 1 and π_{+1} for survey 2.
- The hypothesis of interest is that of *marginal homogeneity*:

$$H_0 : \pi_{1+} = \pi_{+1}$$

- When H_0 holds,

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

Thus, the test of marginal homogeneity is equivalent to testing

$$H_0 : \pi_{12} = \pi_{21},$$

or symmetry across the main diagonal.

- When H_0 holds we would expect about the same frequencies for n_{12} and n_{21} . Let $n^* = n_{12} + n_{21}$ be the total count in the off-diagonal cells. Under H_0 ,

$$n_{12} \sim \text{Binomial}(n^*, 1/2).$$

- When $n^* > 10$, we use the test statistic

$$Z = \frac{n_{12} - \frac{n^*}{2}}{\left[n^* \left(\frac{1}{2}\right) \left(\frac{1}{2}\right)\right]^{1/2}} = \frac{n_{12} - n_{21}}{(n_{12} + n_{21})^{1/2}}$$

- We reject H_0 if $Z^2 > \chi_{1,\alpha}^2$.
- This test is known as **McNemar's test**.
- Let $p_{ij} = n_{ij}/n$ denote the sample proportion in the (i, j) cell. Then a $100(1 - \alpha)\%$ confidence interval for $\pi_{1+} - \pi_{+1}$ is given by

$$p_{1+} - p_{+1} \pm z_{\alpha/2} \hat{se}(p_{1+} - p_{+1})$$

where $\hat{se}(p_{1+} - p_{+1}) = \sqrt{\hat{V}(p_{1+} - p_{+1})}$ and

$$\begin{aligned} \hat{V}(p_{1+} - p_{+1}) &= \frac{p_{1+}(1 - p_{1+}) + p_{+1}(1 - p_{+1}) - 2(p_{11}p_{22} - p_{12}p_{21})}{n} \\ &= \left[(n_{12} + n_{21}) - \frac{(n_{12} - n_{21})^2}{n} \right] / n^2 \end{aligned}$$

8.2 Logistic Regression for Matched Pairs

We now relate the analyses in the previous section to logistic regression models.

8.2.1 Marginal Models for Marginal Proportions

Let (Y_1, Y_2) denote the two responses where “1” denotes category 1 (success) and “0” denotes category 2. In the previous table, Y_1 = the opinion in the first survey and Y_2 = the opinion in the second survey. Then $P(Y_1 = 1) = \pi_{1+}$ and $P(Y_2 = 1) = \pi_{+1}$ are the two marginal probabilities. We estimate these using the marginal proportions p_{1+} and p_{+1} , respectively.

We can form a model using the logit link:

$$\text{logit}[P(Y_1 = 1)] = \alpha + \beta, \quad \text{logit}[P(Y_2 = 1)] = \alpha.$$

This is equivalent to the model

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

where $x_t = 1$ when $t = 1$ and $= 0$ when $t = 2$.

The parameter β is the log of the odds ratio comparing the two marginal distributions. Here $\hat{\beta} = 0.163$ and $e^{0.163} = 1.117 = (944/656)/(880/720)$.

8.2.2 Subject-Specific and Population-Averaged Tables

An alternative way of considering matched-pairs data uses n separate 2×2 partial tables, one for each matched pair. The k^{th} partial table shows the responses (Y_1, Y_2) for the k^{th} subject. The columns represent the two possible outcomes for each response. The following table represents a subject who approved on the first survey and disapproved on the second survey:

	Response	
Survey	Approve	Disapprove
First	1	0
Second	0	1

The full three-way table corresponding to the earlier table would have 1600 partial tables. Of them, 150 would have the above form.

The 1600 subjects provide 3200 observations in a $2 \times 2 \times 1600$ contingency table. If we find the marginal 2×2 table by collapsing over the 1600 subjects, we obtain the following table of marginal totals:

	Response	
Survey	Approve	Disapprove
First	944	656
Second	880	720

The $2 \times 2 \times n$ table with a separate partial table for each of the n matched pairs is called a *subject-specific table*. We will use *conditional models* for its analysis. The original 2×2 table that cross-classifies for all subjects is called a *population-averaged table*. We use marginal models such as the logit model above for the analysis of such tables.

8.2.3 Conditional Logistic Regression for Matched Pairs

A conditional model for the subject-specific tables allows each subject to have their own probability distribution. We have matched observations Y_{it} where $i = 1, \dots, n$ refers to the subject and $t = 1, 2$ to the response. The form of the model is

$$\text{logit}[P(Y_{i1} = 1)] = \alpha_i + \beta, \quad \text{logit}[P(Y_{i2} = 1)] = \alpha_i.$$

or

$$\text{logit}[P(Y_{it} = 1)] = \alpha_i + \beta x_{it} \quad \text{with } x_{i1} = 1 \text{ and } x_{i2} = 0.$$

The probabilities of success for subject i are

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

- The $\{\alpha_i\}$ parameters allow the probabilities to vary among the subjects.
- A large positive value of α_i relative to β results in a high probability of success for each observation.
- A large negative value of α_i relative to β results in a low probability of success for each observation.
- This model assumes for each subject that the odds of success for observation 1 are $\exp(\beta)$ times the odds for success for observation 2.
- Since each partial table refers to a single subject, this conditional association is a subject-specific effect.
- If $\beta = 0$, there is marginal homogeneity and the probability of success is the same for both observations.
- We wish to carry out inference on β to compare the distributions for the two observations. However, the model has an α_i parameter for each subject. This causes problems with the fitting process and the properties of MLEs. To handle this, we use *conditional maximum likelihood* which maximizes the likelihood as a function of β for a conditional distribution that eliminates the subject parameters.

In terms of our original two-way table, the conditional mle of β and its estimated standard error are given by

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

For our example,

$$\hat{\beta} = \log(150/86) = 0.556 \quad \text{and} \quad \hat{se} = \sqrt{1/86 + 1/150} = 0.135.$$

The estimated odds-ratio is $\exp(\hat{\beta}) = n_{12}/n_{21} = 1.744$.

- This estimated OR is identical to that using the Mantel-Haenszel estimate of a common OR for several 2×2 tables.
- An alternative model treats $\{\alpha_i\}$ as *random effects*. The $\{\alpha_i\}$ are assumed to be an unobserved sample from a normal distribution with unspecified mean and variance. Our model is then an example of a *mixed model*. The marginal MLE for this model will result in the same estimate of β .

8.2.4 Logistic Regression for Matched Case-Control Studies

In a case-control study with a binary response Y , each case ($Y = 1$) is matched with a control ($Y = 0$) according to criteria that could affect the response. Subjects are then measured on a binary predictor X .

Example: Diagnoses of Diabetes for MI Case-Control Pairs A study of acute myocardial infarction among Navajo Indians matched 144 victims of MI with 144 individuals free of heart disease. Subjects were asked whether they had diabetes ($x = 0$, no; $x = 1$, yes).

MI Controls	MI Cases		Total
	Diabetes	No diabetes	
Diabetes	9	16	25
No Diabetes	37	82	119
Total	46	98	144

The above table differs from the earlier one in that the levels of X rather than the levels of Y define the rows and columns. In each matched pair, one individual has $Y = 1$ and one has $Y = 0$. The possible pairs appear in the following table:

Diabetes	a		b		c		d	
	Case	Control	Case	Control	Case	Control	Case	Control
Yes	1	1	0	1	1	0	0	0
No	0	0	1	0	0	1	1	1

For the i^{th} subject, we consider the model

$$\text{logit}[P(Y_i = 1)] = \alpha_i + \beta x.$$

In this model, the probabilities refer to distribution of Y given X . However, the retrospective study provides information only about the distribution of X given Y . However, we can estimate $\exp(\beta)$ since it refers to the XY odds-ratio which relates to both conditional distributions.

Example: The conditional ML estimate of the odds ratio $\exp(\beta)$ is

$$\exp(\hat{\beta}) = \frac{n_{21}}{n_{12}} = \frac{37}{16} = 2.31.$$

8.3 Square Tables and Tests of Marginal Homogeneity

We extend the idea of comparing dependent proportions in 2×2 tables to examining $I \times I$ square tables where the rows and columns have the same levels. These tables can result when the same characteristics are recorded on individuals at two times.

The table satisfies *marginal homogeneity* if

$$\pi_{i+} = \pi_{+i} \text{ for all } i.$$

We used McNemar's test for marginal homogeneity of a 2×2 table.

Let (Y_1, Y_2) denote the two responses where “1” denotes category 1 (success) and “0” denotes category 2. In the previous table, Y_1 = the opinion in the first survey and Y_2 = the opinion in the second survey. Then $P(Y_1 = 1) = \pi_{1+}$ and $P(Y_2 = 1) = \pi_{+1}$ are the two marginal probabilities. We formed a logit model for these marginal probabilities:

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

where $x_t = 1$ when $t = 1$ and $= 0$ when $t = 2$.

We can extend the model for binary matched pairs to a baseline-category logit model:

$$\log \left[\frac{P(Y_t = j)}{P(Y_t = I)} \right] = \alpha_j + \beta_j x_t, \quad t = 1, 2, \quad j = 1, \dots, I - 1,$$

where $x_1 = 1$ and $x_2 = 0$. This model has $2(I - 1)$ parameters for the $2(I - 1)$ marginal probabilities, and hence, it is saturated.

Marginal homogeneity is the special case where $\beta_1 = \dots = \beta_{I-1} = 0$. To test marginal homogeneity, fitted values $\{\hat{\mu}_{ij}\}$ are obtained assuming marginal homogeneity and are compared to the observed counts $\{n_{ij}\}$ using G^2 or X^2 . Under the null hypothesis of marginal homogeneity, these statistics have an approximately chi-squared distribution with $df = I - 1$.

An alternative approach due to Bhapkar is based upon the asymptotic normality of the differences of estimated marginal proportions, $p_{+i} - p_{i+}$. The appropriate quadratic form in the differences forms a chi-squared statistic with $df = I - 1$.

Example:

In the 2010 General Social Survey, the respondent's region of residence as well as their region of residence at age 16 were recorded in the following table.

Residence at Age 16	Residence in 2010				Total
	Northeast	Midwest	South	West	
Northeast	266	15	61	28	370
Midwest	10	414	50	40	514
South	8	22	578	22	630
West	7	6	27	301	341
Total	291	457	716	391	1855

Relatively few people changed locations with 84% of the observations falling on the main diagonal. However, the migration was not symmetric resulting in sample marginal percentages of (19.9, 27.7, 34.0, 18.4) at age 16 and (15.7, 24.6, 38.6, 21.2) in 2010.

8.4 Square Tables and Tests of Symmetry

The probabilities in a two-way table satisfy *symmetry* if

$$\pi_{ij} = \pi_{ji} \text{ for all } i, j.$$

If a table satisfies symmetry, then it also satisfies marginal homogeneity. However, a table ($I > 2$) can satisfy marginal homogeneity without satisfying symmetry.

8.4.1 Testing for Symmetry

We can write the condition for symmetry in *logit* form:

$$\log \left(\frac{\pi_{ij}}{\pi_{ji}} \right) = 0 \text{ all } i, j.$$

We can also write the expected cell frequencies as a log-linear model:

$$\log(\mu_{ij}) = \lambda + \lambda_i + \lambda_j + \lambda_{ij}$$

where

$$\lambda_{ij} = \lambda_{ji} \text{ all } i, j.$$

This is a special case of the saturated model where $\lambda_i^X = \lambda_i^Y = \lambda_i$ and $\lambda_{ij}^{XY} = \lambda_{ji}^{XY} = \lambda_{ij}$.

For this model, the expected cell counts have the form

$$\hat{\mu}_{ij} = \hat{\mu}_{ji} = \frac{n_{ij} + n_{ji}}{2} \text{ and } \hat{\mu}_{ii} = n_{ii}.$$

The adjusted residuals equal

$$r_{ij} = \frac{n_{ij} - n_{ji}}{\sqrt{n_{ij} + n_{ji}}}.$$

The chi-squared statistic for symmetry is

$$X^2 = \sum \sum_{i < j} \frac{(n_{ij} - n_{ji})^2}{n_{ij} + n_{ji}}$$

with $df = \frac{I(I-1)}{2}$.

8.4.2 Quasi-Symmetry

The model for symmetry is so simple that it rarely fits well. A generalization of the symmetry model that allows for different marginal distributions is the *quasi-symmetry* model. In this model, the main effect terms are allowed to differ:

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}$$

where

$$\lambda_{ij} = \lambda_{ji}.$$

The symmetry model is a special case where $\lambda_i^X = \lambda_i^Y$.

The quasi-symmetry model can also be written in logit form:

$$\log\left(\frac{\pi_{ij}}{\pi_{ji}}\right) = \beta_i - \beta_j.$$

One parameter is redundant and we set $\beta_I = 0$. The symmetry model is a special case with $\beta_1 = \cdots = \beta_{I-1} = 0$.

We can fit the quasi-symmetry model either as a loglinear model or a logit model.

- For the loglinear model, we define a variable s_{ymm} that takes on a different value for each off-diagonal (n_{ij}, n_{ji}) pair. We then fit a loglinear model with nominal row and s_{ymm} effects.
- For the logit model, we treat each pair of (n_{ij}, n_{ji}) values as an independent binomial variable, ignoring the diagonal entries. Set up I explanatory variables corresponding to the coefficients of the $\{\beta_i\}$ variables. For $\log(\pi_{ij}/\pi_{ji})$, this variable for β_i is 1, for β_j is -1 , and for other variables is zero.

We can use the deviances of the symmetry model and the quasi-symmetry model to test whether the symmetry model is adequate:

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

We use $I - 1$ degrees of freedom. This test can also be viewed as a test of marginal homogeneity for the quasi-symmetry model.

Example: Premarital and Extramarital Sex

The following table is taken from the 1989 General Social Survey. The response categories are

- 1 = always wrong
- 2 = almost always wrong
- 3 = wrong only sometimes
- 4 = not wrong at all

Premarital	Extramarital Sex				
Sex	1	2	3	4	Total
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

The symmetry model does not fit the data well with $G^2 = 402.2$ and $X^2 = 297.6$ with 6 *d.f.* On the other hand, the quasi-symmetry model does fit the model well with $G^2 = 1.3646$ and $X^2 = 0.8683$ with 3 *d.f.* The likelihood ratio test comparing the two models results in

$$G^2(S|QS) = G^2(S) - G^2(QS) = 402.2 - 1.4 = 400.8$$

with 3 *d.f.* This is strong indication of the inadequacy of the symmetry model.

8.5 Ordinal Quasi-Symmetry Model

When the row and column categories are ordinal, we can obtain more parsimonious models for quasi-symmetry by taking into account the ordering. We let $u_1 \leq u_2 \leq \dots \leq u_I$ be the scores for rows and columns. The ordinal quasi-symmetry loglinear model is given by

$$\log \mu_{ij} = \lambda + \lambda_i + \lambda_j + \beta u_j + \lambda_{ij}$$

with $\lambda_{ij} = \lambda_{ji}$. For this model, the difference in the main effects for row and columns for j is a constant multiple of the score for level j . If $\beta > 0$, $P(X \leq j) > P(Y \leq j)$ for $j < I$. This says that lower values are more likely for rows than for columns. If $\beta < 0$, the opposite pattern occurs. When $\beta = 0$, the model corresponds to the symmetry model.

The ordinal quasi-symmetry model also has a logit representation,

$$\log \left(\frac{\pi_{ij}}{\pi_{ji}} \right) = \beta(u_j - u_i).$$

When $\beta = 0$, the model corresponds to the symmetry model. To estimate β in this model, we can identify n_{ij} as the number of successes in $n_{ij} + n_{ji}$ trials and fit a logit model with no intercept and predictor equal to $u_i - u_j$.

The symmetry model does not fit the data well with $G^2 = 402.2$ and $X^2 = 297.6$ with 6 *d.f.* On the other hand, the ordinal quasi-symmetry model does fit the model well with $G^2 = 2.0972$ and $X^2 = 2.0844$ with 5 *d.f.* The likelihood ratio test comparing the two models results in

$$G^2(S|OQS) = G^2(S) - G^2(OQS) = 402.2 - 2.1 = 400.1$$

with 1 *d.f.* This is strong indication of the inadequacy of the symmetry model.

8.6 Analyzing Rater Agreement

One use of square tables is to analyze the agreement between two observers. We represent the ratings due to one observer using the rows of the table and the ratings of a second observer using the columns of the table. Each cell represents one particular combination of the observers' ratings. The diagonal cells represent the cases where the observers agree. The probability that both raters classify a subject in category i is π_{ii} . The total probability of agreement is the sum $\sum_i \pi_{ii}$. We are interested in developing measures of agreement and detecting patterns of disagreement.

Example: Two pathologists labelled X and Y separately classified 118 slides on the presence and extent of carcinoma of the uterine cervix using 4 ordered categories (see Agresti, p. 260). The following table contains the resulting data:

Pathologist	Pathologist Y			
X	1	2	3	4
1	22	2	2	0
2	5	7	14	0
3	0	2	36	0
4	0	1	17	10

One way to assess the degree of agreement in the table is to compare the cell counts $\{n_{ij}\}$ to the values that are predicted by the independence model. The values in the following table are the standardized Pearson residuals:

Pathologist	Pathologist Y			
X	1	2	3	4
1	8.5	-0.5	-6.0	-1.8
2	-0.5	3.2	-0.5	-1.8
3	-4.1	-1.2	5.5	-2.3
4	-3.3	-1.3	0.3	5.9

The large positive residuals on the diagonal indicate that there is much greater agreement between the raters than what would be predicted by the independence model. The independence model fits the data poorly ($G^2 = 118.0$, $df = 9$).

8.6.1 Quasi-Independence Model

Often in square tables, particularly in rater agreement tables, there is much agreement in the responses resulting in much higher counts on the diagonal than one would expect in the independence model. A model that adds a term to take this into account is the *quasi-independence model*:

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \delta_i I(i = j).$$

Here $I(i = j) = 1$ if $i = j$ and $= 0$ otherwise. This model treats the diagonal cells differently than the other cells.

- When $\delta_i > 0$, there are more agreements for outcome i than would be expected under independence.
- For the remaining cells, the independence model still applies. This says that if someone changed their opinion, the response the second time is independent of the response the first time.

There are I more parameters than the independence model, so the quasi-independence model has $df = (I - 1)^2 - I$.

The quasi-independence model was fit to the pathologist data resulting in $G^2 = 13.2$, $df = 5$ indicating better fit than the independence model but still some lack of fit.

The quasi-symmetry model, which also fits the main diagonal perfectly, is more complex than the quasi-independence model and allows some association off the main diagonal. The quasi-symmetry model was fit to the pathologist data resulting in $G^2 = 1.0$, $df = 2$ indicating a very good fit to the data. The following table presents the observed values together with the fitted values for the **quasi-independence/ quasi-symmetry models**:

Pathologist X	Pathologist Y			
	1	2	3	4
1	22	2	2	0
	22	0.7/2.6	3.3/1.6	0.0/0.0
2	5	7	14	0
	2.4/4.6	7	16.6/14.4	0.0/0.0
3	0	2	36	0
	0.8/0.4	1.2/1.6	36	0.0/0.0
4	0	1	17	10
	1.9/0.0	3.0/1.0	13.1/17.0	10

8.6.2 Kappa Measure of Agreement

Cohen's kappa is the most commonly used measure of agreement between raters in a square table. The measure is based on comparing the observed agreement in the table to what would be expected if the ratings were independent. The probability of agreement is $\sum_i \pi_{ii}$. If the ratings were independent, then $\pi_{ii} = \pi_{i+}\pi_{+i}$. The excess probability of agreement over independence is

$$\sum_i \pi_{ii} - \sum_i \pi_{i+}\pi_{+i}.$$

This value is scaled using the maximum probability of agreement of 1 to obtain Cohen's kappa:

$$\kappa = \frac{\sum_i \pi_{ii} - \sum_i \pi_{i+}\pi_{+i}}{1 - \sum_i \pi_{i+}\pi_{+i}}.$$

The sample version $\hat{\kappa}$ is obtained by substituting the sample proportions in the above expression. For the pathologist table, we can obtain $\hat{\kappa} = 0.493$ indicating that the difference between the observed agreement and that expected under independence is about $1/2$ the maximum possible difference.

We note the κ treats the categories as nominal. When the categories are ordinal, larger differences between ratings are more serious and a *weighted kappa measure* may be used.

8.7 Bradley-Terry Model for Paired Comparisons

Sometimes categorical outcomes result from pairwise comparisons. In some cases it would be difficult to rank I categories, but easier to compare category i to category j for any pair of categories i and j . The fitted model will provide a ranking of the categories in addition to estimates of the probability that any category i would be rated higher than category j .

- This model is often applied to athlete contests where the categories represent teams or individuals. The pairwise comparison is made using competitions where you record whether i or j was victorious.
- The model can be compared to the comparison of a number of brands of a product. The judge records whether brand i is preferred to brand j or brand j is preferred to brand i .

The Bradley-Terry model is a logit model for paired preference data. Let Π_{ij} represent the probability that category i is “the winner” over category j in the competition between i and j . The probability the j is the winner is $\Pi_{ji} = 1 - \Pi_{ij}$. The Bradley-Terry model has category parameters $\{\beta_i\}$ such that

$$\text{logit}(\Pi_{ij}) = \log(\Pi_{ij}/\Pi_{ji}) = \beta_i - \beta_j.$$

One parameter is redundant and a constraint must be used by software. The probability of category i winning equals $\frac{1}{2}$ when $\beta_i = \beta_j$ and is greater than $\frac{1}{2}$ when $\beta_i > \beta_j$.

This logit model is equivalent to the quasi-symmetry model. We can use the fitted quasi-symmetry model to obtain estimates of the parameters $\{\beta_i\}$. In fact, it can be shown that

$$\beta_i = \lambda_i^X - \lambda_i^Y,$$

where λ_i^X and λ_i^Y are the parameters from the quasi-symmetry model. We can then use $\{\beta_i\}$ to compute the probability that i defeats j :

$$\Pi_{ij} = \frac{\exp(\beta_i)}{\exp(\beta_i) + \exp(\beta_j)}$$

Example: Ranking the teams in the NBA Western Conference.

There are eight teams that qualified for the playoffs in the Western Division of the NBA. During the 2008-09 regular season, these eight teams played each other either three or four times. The following table portrays the results of their games:

Statistics 659

Winner	Loser								Total
	Dallas	Denver	Houston	LA	NO	Portland	SA	Utah	
Dallas	—	0	2	0	1	3	2	2	10
Denver	4	—	1	1	2	2	2	2	14
Houston	2	3	—	0	3	2	2	2	14
Los Angeles	3	3	4	—	3	2	2	2	19
New Orleans	3	2	1	1	—	2	2	0	11
Portland	0	2	1	2	2	—	3	2	12
San Antonio	2	1	2	1	2	1	—	3	12
Utah	2	2	2	1	3	2	0	—	12
Total	16	13	13	6	16	14	13	13	104

The Bradley-Terry model was fit to these data. The fit was adequate ($G^2 = 26.6$, $df = 21$). Los Angeles was ranked highest by far with $\hat{\beta} = 1.12$ with the others ranging from -0.36 to 0.16 . However, the standard errors of the estimates were large making strong conclusions impossible for this small data set.