Biometrical Letters
Vol. 43 (2006), No. 2, 91-98

Linear column-parameter symmetry model for square contingency tables: application to decayed teeth data

Sadao Tomizawa*, Nobuko Miyamoto, Mari Iwamoto

Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science, Noda City, Chiba, 278-8510, Japan
*e-mail: tomizawa@is.noda.tus.ac.jp

SUMMARY

For square contingency tables with ordered categories, this paper proposes a new asymmetry model which indicates that the odds for two symmetric cell probabilities is an exponential function of column values. Using this model, Japanese decayed teeth data are analyzed, and it is shown that the right (upper) teeth are worse than the left (lower) teeth.

Key words: Conditional symmetry, Decayed teeth, Model, Ordinal data, Square contingency table, Symmetry.

1. Introduction

For an r x r square contingency table with the same ordinal row and column classifications, let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1,\dots,r; j = 1,\dots,r$).

Goodman (1979) considered the diagonals-parameter symmetry (DPS) model defined by

$$p_{ij} = \begin{cases} \delta_{j-i}\psi_{ij} & (i < j), \\ \psi_{ij} & (i \geq j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$. This model may be expressed as

$$\frac{p_{ij}}{p_{ji}} = \delta_{j-i} \quad (i < j).$$

This indicates that the odds that an observation will fall in cell (i, j), instead of cell (j, i), $i < j$, depends on only the distance $j - i$ from the main diagonal of the table. So, the probability that an observation will fall in cell (i, j), $i < j$, is δ_{j-i} times the probability that the observation falls in cell (j, i).

Special cases of the DPS model obtained by putting $\delta_1 = \ldots = \delta_{r-1} = 1$ and $\delta_1 = \ldots = \delta_{r-1}$ $(= \delta)$ are the usual symmetry (S) model (Bowker, 1948; Bishop, Fienberg and Holland, 1975, p. 282) and the conditional symmetry (CS) model (McCullagh, 1978), respectively (also see Tomizawa, 1993). Thus, the S model indicates that the probability that an observation will fall in cell (i, j), $i < j$, equals the probability that the observation falls in cell (j, i). The CS model indicates that the probability that an observation will fall in cell (i, j), $i < j$, is δ times the probability that the observation falls in cell (j, i).

Consider the data in Table 1. Table 1a (Table 1b) is constructed from the data of the decayed teeth of 349 men (363 women) aged 18-39, for the patients visiting a dental clinic in Sapporo City, Japan, from 2001 to 2005. Tables 1a and 1b are classified by the numbers of decayed teeth in the left side of the mouth of a patient and those in the right side. Note that each of these patients has at least one decayed tooth. Tables 1c and 1d are re-classified by the numbers of decayed teeth in the lower side of the mouth of a patient and those in the upper side.

Table 1. Decayed teeth data of 349 men and 363 women aged 18-39, for patients visiting a dental clinic in Sapporo City, Japan, from 2001 to 2005. (The parenthesized values are the MLEs of expected frequencies under the LCPS model)

(a) For men with left and right decayed teeth

Left (numbers of Decayed teeth)		Right (numbers of decayed teeth) 0-4 (1)	5-8 (2)	9+ (3)	Total
0-4	(1)	118	37	2	157
		(118.00)	(34.68)	(2.75)	
5-8	(2)	21	87	23	131
		(23.33)	(87.00)	(23.41)	
9+	(3)	2	11	48	61
		(1.25)	(10.59)	(48.00)	
Total		141	135	73	349

(b) For women with left and right decayed teeth

Left	Right (1)	(2)	(3)	Total
(1)	103	45	1	149
	(103.00)	(45.00)	(2.51)	
(2)	35	84	33	152
	(34.80)	(84.00)	(31.39)	
(3)	3	17	42	62
	(1.49)	(18.61)	(42.00)	
Total	141	146	76	363

(c) For men with lower and upper decayed teeth

Lower	Upper			Total
	(1)	(2)	(3)	
(1)	115	55	25	195
	(115.00)	(54.54)	(23.83)	
(2)	16	49	60	125
	(16.46)	(49.00)	(61.40)	
(3)	1	7	21	29
	(2.17)	(5.60)	(21.00)	
Total	132	111	106	349

(d) For women with lower and upper decayed data

Lower	Upper			Total
	(1)	(2)	(3)	
(1)	97	62	15	174
	(97.00)	(62.78)	(15.54)	
(2)	20	63	75	158
	(19.22)	(63.00)	(74.06)	
(3)	2	6	23	31
	(1.46)	(6.94)	(23.00)	
Total	119	131	113	363

For each data set, let $\hat{p}_{ij} = n_{ij} / n$, where n = $\Sigma\Sigma n_{ij}$ and n_{ij} is the observed frequency in the (i, j)th cell of the table. Table 2 gives the values of $\{\beta_{ij} / \beta_{ji}\}$, $i<j$, for the data in Table 1.

Table 2. For Table 1, the values of $\{ \hat{p}_{ij} / \hat{p}_{ji} \}, i < j$

(a) For Table 1a

	J=2	3
i=1	1.76	1.00
2	-	2.09

(b) For Table 1b

	J=2	3
i=1	1.29	0.33
2	-	1.94

(c) For Table 1c

	J=2	3
i=1	3.44	25.00
2	-	8.57

(d) For Table 1d

	j=2	3
i=1	3.10	7.50
2	-	12.50

From Table 2a, it is likely that for the data in Table 1a, the values of odds $\{p_{ij}/p_{ji}\}$, $i<j$, are constant. Also, from Table 2b, it is likely that for the data in Table 1b, the values of odds $\{p_{ij}/p_{ji}\}$, $i<j$, may be constant or may depend on only the distance $j - i$ from the main diagonal of the table. However, from Tables 2c and 2d, it is unlikely that for the data in Tables 1c and 1d, the values of odds are constant or depend on the distance $j - i$. Namely, for the data in Tables 1c and 1d, it is unlikely that each of the S, CS and DPS models fits well (see Table 3). We see now from Tables 2c and 2d that for the data in Tables 1c and 1d, the values of odds $\{\hat{p}_{ij}/\hat{p}_{ji}\}$, $i<j$, are greater when $j = 3$ than when $j = 2$, especially for the women data in Table 1d, the values of odds when $j = 3$ are close to the square of the value of odds when $j = 2$. So, instead of the S, CS and DPS models, we are now interested in fitting a new model, in which the odds $\{p_{ij}/p_{ji}\}$, $i<j$, are an exponential function of column j.

Table 3. Values of likelihood ratio statistic G^2 applied to the data in Table 1

			G^2		
models	df	Table 1a	Table 1b	Table 1c	Table 1d
S	3	8.80*	7.51**	98.24*	103.33*
CS	2	0.51	3.19	7.44*	9.40*
DPS	1	0.14	1.23	3.72**	9.18*
LCPS	2	1.03	2.59	1.22	0.39

* - means significant at the 0.05 level

** - means that the value is almost at the 5 percent point

The purpose of this paper is (1) to propose such a new model which indicates that the odds are an exponential function of column values, and (2) to analyze the data in Table 1 using the models.

2. A new model

Consider a new model defined by

$$p_{ij} = \begin{cases} \theta^{j-1}\Psi_{ij} & (i < j), \\ \Psi_{ij} & (i \geq j), \end{cases} \qquad (2.1)$$

where $\Psi_{ij} = \Psi_{ji}$. This model states that the probability that an observation will fall in cell (i, j), $i < j$, is θ^{j-1} times the probability that the observation falls in cell (j, i). This model may be expressed as

$$\frac{p_{ij}}{p_{ji}} = \theta^{j-1} \quad (i < j).$$

Thus, this says that the odds that an observation will fall in cell (i, j), instead of cell (j, i), $i < j$, is θ^{j-1}. Namely, the odds are an exponential function of column values j. The parameter θ is the odds that an observation will fall in cell (1,2) instead of cell (2,1). We shall refer to (2.1) as the linear column-parameter symmetry (LCPS) model.

Let X and Y denote the row and column variables, respectively. Under the LCPS model, $\theta > 1$ is equivalent to $P(X < Y) > P(X > Y)$. Therefore the parameter θ in the LCPS model would be useful for making inferences such as X tends to be less (rather than greater) than Y or vice versa, according as whether θ is greater or less than 1.

In addition, under the LCPS model, $\theta > 1$ is equivalent to $P(X \leq i) > P(Y \leq i)$ for $i = 1, \ldots, r-1$. Therefore the parameter θ in the LCPS model would be useful for making inferences such as X is stochastically less than Y or vice versa.

3. Goodness-of-fit test

Assume that a multinomial distribution is applied to the $r \times r$ table. The maximum likelihood estimates (MLEs) of expected frequencies under the LCPS model could be obtained using an iterative procedure, for example, the general iterative procedure for log-linear model of Darroch and Ratcliff (1972), or the Newton-Raphson method in the log-likelihood equations.

The likelihood ratio statistic for testing the goodness-of-fit of the model is

$$G^2 = 2\sum_i \sum_j n_{ij} \log\left(\frac{n_{ij}}{\hat{m}_{ij}}\right),$$

where $\hat{m}_{ij}$ is the MLE of expected frequency m_{ij} under the model. The number of degrees of freedom (df) for testing the goodness-of-fit of the LCPS model is $(r + 1)(r - 2)/2$, which is equal to that for the CS model. The number of df for the DPS model is $(r - 2)(r - 1)/2$.

4. Analysis of decayed teeth data

4.1 Analysis of Table 1a

Consider the data in Table 1a. We see from Table 3 that the S model fits these data poorly, however, the CS, DPS, and LCPS models fit these data well. For testing that the CS model holds (i.e., $\delta_1 = \delta_2$) assuming that the DPS model holds, the difference between the likelihood ratio statistics is 0.37 with 1 df. Thus the CS model is preferable to the DPS model. In addition, the value of G^2 for the CS model is less than that for the LCPS model with both 2 df. Thus the CS model may be preferable to the LCPS model for these data.

Under the CS model, the MLE of δ is $\hat{\delta} = 1.824$. Hence, under this model, the probability that a man's left decayed teeth grade is i (i =1,2) and his right decayed teeth grade is j ($> i$) is estimated to be $\hat{\delta} = 1.824$ times the probability that the man's right decayed teeth grade is i and his left decayed teeth grade is j, Since $\hat{\delta} > 1$, under this model, a man's right teeth are estimated to be worse than his left teeth.

We note that under the LCPS model, the MLE of θ is $\hat{\theta} = 1.487$ ($\hat{\theta}^2 = 2.210$), and so we conclude again that a man's right teeth are estimated to be worse than his left teeth.

4.2 Analysis of Table 1b

Consider the data in Table 1b. We see from Table 3 that the S model does not fit these data so well, however, the CS, DPS, and LCPS models fit these data well. From the test based on the difference between the values of G^2, the CS model would be preferable to the DPS model. In addition, the value of G^2 for the LCPS model is less than that for the CS model with both 2 df, Thus the LCPS model may be preferable to the CS model for these data.

Under the LCPS model, the MLE of θ is $\hat{\theta} = 1.299$ ($\hat{\theta}^2 = 1.687$). Hence, under this model, (1) the probability that a woman's left decayed teeth grade is 1 and her right decayed teeth grade is 2 is estimated to be $\hat{\theta} = 1.299$ times the probability that the woman's right decayed teeth grade is 1 and her left decayed teeth grade is 2, and (2) the probability that a woman's left decayed teeth grade is i (i =1,2) and her right decayed teeth grade is 3 is estimated to be $\hat{\theta}^2 = 1.687$ times the probability that the woman's right decayed teeth grade is i and her left decayed teeth grade is 3. Since $\hat{\theta} > 1$, under this model, a woman's right teeth are estimated to be worse than her left teeth.

4.3 Analysis of Table 1c

Consider the data in Table 1c. We see from Table 3 that the S and CS models fit these data poorly, and the DPS model does not fit these data so well, however, the LCPS model fits these data well.

Under the LCPS model, the MLE of θ is $\hat{\theta} = 3.312$ ($\hat{\theta}^2 = 10.972$). Hence, under this model, (1) the probability that a man's lower decayed teeth grade is 1 and his upper decayed teeth grade is 2 is estimated to be $\hat{\theta} = 3.312$ times the probability that the man's upper decayed teeth grade is 1 and his lower decayed teeth grade is 2, and (2) the probability that a man's lower decayed teeth grade is i ($i = 1,2$) and his upper decayed teeth grade is 3 is estimated to be $\hat{\theta}^2 = 10.972$ times the probability that the man's upper decayed teeth grade is i and his lower decayed teeth grade is 3. Since $\hat{\theta} > 1$, under this model, a man's upper teeth are estimated to be worse than his lower teeth.

4.4 Analysis of Table 1d

Consider the data in Table 1d, We see from Table 3 that the S, CS, and DPS models fit these data poorly, however, the LCPS model fits these data very well.

Under the LCPS model, the MLE of θ is $\hat{\theta} = 3.267$ ($\hat{\theta}^2 = 10.677$). Hence, under this model, (1) the probability that a woman's lower decayed teeth grade is 1 and her upper decayed teeth grade is 2 is estimated to be $\hat{\theta} = 3.267$ times the probability that the woman's upper decayed teeth grade is 1 and her lower decayed teeth grade is 2, and (2) the probability that a woman's lower decayed teeth grade is i ($i = 1,2$) and her upper decayed teeth grade is 3 is estimated to be $\hat{\theta}^2 = 10.677$ times the probability that the woman's upper decayed teeth grade is i and her lower decayed teeth grade is 3. Since $\hat{\theta} > 1$, under this model, a woman's upper teeth are estimated to be worse than her lower teeth.

5. Concluding remarks

In Section 2 we proposed a new model, i.e., the LCPS model, which indicates that the odds for two symmetric cell probabilities is an exponential function of column values (i.e. which indicates the structure of asymmetry of cell probabilities). The parameter θ in this model would be useful for making inferences such as that the row variable X is stochastically less than the column variable Y or vice versa.

By applying the LCPS model, we analyzed the decayed teeth data, and we concluded that for both men's and women's data (1) the right teeth are estimated to be worse than the left teeth, and (2) the upper teeth are estimated to be worse than the lower teeth. In addition, we emphasize that the difference between the degrees of the upper and the lower decayed teeth are surprisingly very great (from the values of $\hat{\theta}$ and $\hat{\theta}^2$ obtained under the LCPS model, see Sections 4.3 and 4.4).

We comment that many Japanese tend to clean their own teeth with the toothbrush in the right hand, and so may not be able to wash the right teeth well, though we have no sufficient scientific evidence.

Finally we note that Tomizawa, Miyamoto and Hatanaka (2001) considered a measure (instead of a model) for representing the degree of departure from the S model, however, the approach using the measure is entirely different from the approach of model-fitting and the measure cannot infer the structure of asymmetry of probabilities shown by the models (e.g. CS, DPS and LCPS models). Hence, the interpretations (obtained by models) for decayed teeth data shown in Section 4 cannot be obtained using the measure.

Acknowledgements

The authors would like to thank the referee for the helpful comments.

References

Bishop Y.M.M., Fienberg S.E., Holland P.W. (1975). Discrete Multivariate Analysis: Theory and Practice. Cambridge, Massachusetts, The MIT Press.

Bowker A.H. (1948). A test for symmetry in contingency tables. Journal of the American Statistical Association 43, 572-574.

Darroch J.N., Ratcliff D. (1972). Generalized iterative scaling for log-linear models. Annals of Mathematical Statistics 43, 1470-1480.

Goodman L.A. (1979). Multiplicative models for square contingency tables with ordered categories. Biometrika 66, 413-418.

McCullagh P. (1978). A class of parametric models for the analysis of square contingency tables with ordered categories. Biometrika 65, 413-418.

Tomizawa S. (1993). Diagonals-parameter symmetry model for cumulative probabilities in square contingency tables with ordered categories. Biometrics 49, 883-887.

Tomizawa S., Miyamoto N., Hatanaka Y. (2001). Measure of asymmetry for square contingency tables having ordered categories. Australian and New Zealand Journal of Statistics 43, 335-349.