数理解析研究所講究録
第 247 巻 1975 年 83-105

A LOGISTIC REGRESSION ANALYSIS OF BIVARIATE BINARY DATA: A METHOD OF ASSESSING THE ASSOCIATION BETWEEN A PAIR OF BINARY RESPONSES

By

K. Ito and A. Kudô

Nanzan University and Kyushu University

1. Introduction

Given a set or sets of multivariate data which are to be interpreted by means of regression analysis, we proposed to carry out a simultaneous analysis of all linear models plausible in the circumstances where data were collected and then determine an adequate model, if any, from among the set of models, keeping in mind the hierarchy among them. (cf. [5], [6])

For the purpose of demonstration, we use the data in Table 1, where 70 male and 115 female graduates from a certain university are summerized in the 2×2×3×3 tables. The four rows are in terms of performances in general education (Y_1=1 or 0) and major field (Y_2=1 or 0). The nine columns are in terms of high school (H.S.) grades and college entrance examination (C.E.E.) grades, both of which are of three levels: A, B and C.

The main concern in this paper is to demonstrate the methodological and/or logical aspects in determining an adequate model, with the hypothesis in mind that Y_1 and Y_2 are independent.

In Section 2, we shall present a schema of analysis by using a tetranomial logistic model. We shall also analyse the data assuming that Y_1 and Y_2 are independent. As the outcomes of the analysis assuming independence do not fit our data, we shall investigate, in Section 3, how Y_1 and Y_2 are associated. So far as the present authors are aware, there does not seem to exist any serious trial for investigating the structure of association between two binary responses. Our investigation will be of some interest in this respect. Section 4 is devoted to the discussions in the logical aspects involved in this analysis.

2. Logistic Regression Analysis

Let the individual graduates be associated with two binary variables

$$Y_1 = \begin{cases} 1 & \text{if performances are excellent in general education,} \\ 0 & \text{otherwise,} \end{cases}$$

$$Y_2 = \begin{cases} 1 & \text{if performances are excellent in major field,} \\ 0 & \text{otherwise,} \end{cases}$$

and as independent variables in the regression model, the following four are defined

$$x_1 = 1 \text{ (A dummy variable taking constant 1)}$$

$$x_2 = \begin{cases} 0 & \text{if male,} \\ 1 & \text{if female,} \end{cases}$$

$$x_3 = \begin{cases} 1 & \text{if high school grades A,} \\ 0 & \text{if " " " B,} \\ -1 & \text{if " " " C,} \end{cases}$$

$$x_4 = \begin{cases} 1 & \text{if C. E. E. grades A,} \\ 0 & \text{if " " " " B,} \\ -1 & \text{if " " " " C.} \end{cases}$$

And let

$Pr[Y_1=0,\ Y_2=0 \mid x_2,\ x_3,\ x_4] = P_0,$

$Pr[Y_1=0,\ Y_2=1 \mid x_2,\ x_3,\ x_4] = P_1,$

$Pr[Y_1=1,\ Y_2=0 \mid x_2,\ x_3,\ x_4] = P_2,$

$Pr[Y_1=1,\ Y_2=1 \mid x_2,\ x_3,\ x_4] = P_3.$

To regress P_a, a=0, 1, 2, 3, on x_2, x_3 and x_4 we shall follow Bock [1], Cox [2], [3], Mantel [7], Mantel & Brown [8] and Ito & Kudô [5] and use a tetranomial logistic linear model:

$$\begin{cases} \lambda_1 = \log P_1/P_0 = \sum_{u=1}^{4} \beta_{1u}\, x_u, \\ \lambda_2 = \log P_2/P_0 = \sum_{u=1}^{4} \beta_{2u}\, x_u, \\ \lambda_3 = \log P_3/P_0 = \sum_{u=1}^{4} \beta_{3u}\, x_u, \end{cases} \qquad (2.1)$$

or in terms of vector notation,

$$\underset{\sim}{\lambda} = \sum_{u=1}^{4} x_u\, \underset{\sim}{\beta}_u\ . \qquad (2.2)$$

The vectors $\underset{\sim}{\beta}_2, \underset{\sim}{\beta}_3$ and $\underset{\sim}{\beta}_4$ may be interpreted as the effects of sex, high school grades and college entrance examination grades, respectively. As regards to the existence or non-existence of these effects, or equivalently being nonzero or zero of these vectors, there are $2^3=8$ possibilities. This gives rise to the notion of a hierarchy of the models.

In order to express the model in our hierarchy, we shall use the notation $(a_1\ a_2\ a_3\ a_4)$, where $a_i=0$ or 1 (i=1,2,3,4) with $\underset{\sim}{\beta}_i=\underset{\sim}{0}$ for $a_i=0$ and $\underset{\sim}{\beta}_i \neq \underset{\sim}{0}$ otherwise. By the very nature of dummy variable x_1, $a_1=1$. These 8 models are shown in the left two columns of Tables 2 and 3.

With (2.1) as the least simple model indicated as (1 1 1 1)

$$\begin{cases} \lambda_1 = \log P_1/P_0 = \beta_{11} x_1, \\ \lambda_2 = \log P_2/P_0 = \beta_{21} x_1, \\ \lambda_3 = \log P_3/P_0 = \beta_{31} x_1, \end{cases} \quad (2.3)$$

is the simplest model indicated as (1 0 0 0).

Now returning to Table 1, there are six nonzero colums in male and eight in female, and these nonzero colums are subscripted by i=1,2,..., 14. We analyse these data as fourteen tetranomial data. Let n_i be the sum of the i-th nonzero column and let

r_{0i}	graduates	with	Y_1=	0	and	Y_2	=0,
r_{1i}	"	"	"	0	and	"	1,
r_{2i}	"	"	"	1	and	"	0,
r_{3i}	"	"	"	1	and	"	1,

with $n_i = \sum r_{ai}$. Then $(r_{0i}, r_{1i}, r_{2i}, r_{3i})$ follows a tetranomial distribution with the parameters n_i and $P_{0i}, P_{1i}, P_{2i}, P_{3i}$.

The log-likelihood function of the parameters in (2.1) or the model (1 1 1 1) in Table 2 is given by

$$L = \text{constant} + \sum_u t_{1u}\beta_{1u} + \sum_u t_{2u}\beta_{2u} + \sum_u t_{3u}\beta_{3u}$$

$$- \sum_i n_i \log\{1 + \exp(\sum_u \beta_{1u}x_{iu}) + \exp(\sum_u \beta_{2u}x_{iu})$$

$$+ \exp(\sum_u \beta_{3u}x_{iu})\} \quad (2.4)$$

where

$$t_{au} = \sum_i r_{ai} x_{iu}, \quad a=1,2,3; \; u=1,2,3,4. \qquad (2.5)$$

The log-likelihood function (2.4) shows that the set of statistics { t_{au}, a=1,2,3; u=1,2,3,4 } forms a set of sufficient statistics for the model (1 1 1 1) involving twelve parameters {β_{au}, a=1,2,3; u=1,2,3,4} .

The log-likelihood function (and the sufficient statistics) in the models simpler than (2.1) in the present hierarchy are given merely by deleting appropriate terms in each of the summations in (2.4) (and in (2.5)). Because of the nature of our hierarchy, the terms with u=1 always remain in each of the summations.

When it is assumed that the variables Y_1 and Y_2 are statistically independent to each other, then

$$P_0 P_3 = P_1 P_2 \qquad (2.6)$$

and in the logistic linear model (2.1), the number of parameters is reduced by the following relations,

$$\beta_{1u} + \beta_{2u} = {}_{3u} \qquad (u=1,2,3,4) \qquad (2.7)$$

Under the assumption of independence, the log-likelihood function of (2.4) is reduced to the one involving eight parameters { β_{au} ; a=1,2 u=1,2,3,4 }

$$L_I = \text{constant} + \sum_u t_{1u}\beta_{1u} + \sum_u t_{2u}\beta_{2u} + \sum_u t_{3u}(\beta_{1u} + \beta_{2u})$$

$$- \sum_i n_i \log \{ 1 + \exp (\sum_u \beta_{1u} x_{iu}) + \exp (\sum_u \beta_{2u} x_{iu})$$

$$+ \exp (\sum_u (\beta_{1u} + \beta_{2u}) x_{iu}) \} \qquad (2.8)$$

The sufficient statistics are reduced correspondingly. The log-likelihood function (and the sufficient statistics) in the models simpler than (2.1) in our hierarchy are obtained in the same manner as in the case of a general tetranomial model.

The maximum likelihood estimates of the unknown parameters in the models in our hierarchy with and without the assumption of independence can be obtained by the usual iterative method of Newton-Raphson. The goodness of fit of each of the models can be tested by the usual χ^2 statistic, where the degree of freedom is, in our present case, 3×14 -(number of paremeters estimated). [9]

The algorithm for computing the maximum likelihood estimates are essentially common all through the models in our hierarchy. Only difference lies between those with the assumption of independence and those without it. This feature is evident in the formulae of the first and the second derivatives of log-likelihood functions presented in the appendix.

The outcomes of fitting the 8 models to the data without and with the assumption of independence are shown in Tables 2 and 3, respectively, in terms of goodness of fit χ^2 and log-likelihood. From these tables it is noted that in the hierarchy of models, no model is considered adequately fitted to the data under the assumption of independence, but model (1 0 1 0) is an adequate one under no assumption of independence, whose maximum likelihood estimate is given by

$$\begin{cases} \hat{\lambda}_1 = \log \hat{P}_1/\hat{P}_0 = -2.107 + 0.665\ x_3, \\ \hat{\lambda}_2 = \log \hat{P}_2/\hat{P}_0 = -1.712 + 2.278\ x_3, \\ \hat{\lambda}_3 = \log \hat{P}_3/\hat{P}_0 = -2.182 + 1.307\ x_3. \end{cases} \tag{2.9}$$

Thus it may be concluded from (2.9) that the effect of high school grades upon a pair of binary-responses Y_1 and Y_2 is significantly exhibited, while no significant effects of sex and C.E.E. grades are present. Table 2 also shows that in a hierarchy where the model (1 0 1 0) is the pivotal model, no less simple models exhibit significant decrease in χ^2. As to the determination procedure of an adequate model, we will discuss it again in Section 4.

3. Association between a pair of binary random variables

The analysis in the previous section indicates that no model is found to fit adequately to the data under the assumption of independence. We therefore would like to know how the components of a pair of binary random variables Y_1 and Y_2 are related to each other.

Let

$$\begin{aligned} p^{(0)} &= \Pr [\ Y_2 = 1 \mid Y_1 = 0] \\ p^{(1)} &= \Pr [\ Y_2 = 1 \mid Y_1 = 1] \end{aligned} \tag{3.1}$$

and corresponding to the pair of probabilities, we consider a pair of logistic linear models

$$\begin{cases} \lambda^{(0)} = \log P^{(0)} / (1 - P^{(0)}) = \sum_u \beta_u^{(0)} x_u , \\ \\ \lambda^{(1)} = \log P^{(1)} / (1 - P^{(1)}) = \sum_u \beta_u^{(1)} x_u . \end{cases} \quad (3.2)$$

The data in Table 1 are interpreted, in this section, as two series of binomial data. The first two rows correspond to the probability $P^{(1)}$ and the last two rows to the probability $P^{(0)}$, and these two probabilities are dependent on the variables x_1, x_2, x_3, and x_4, which were defined in the previous section.

Following [5], we shall interpret the association by means of a combined hierarchy of all pairs of logistic linear models associated with (3.1) and (3.2). This hierarchy is shown in Table 4, where the models are denoted by a quadruple of integers (a_1 a_2 a_3 a_4), with a_i=0,1,2, where a_i=0 (and =1) mean $\beta_i^{(0)} = \beta_i^{(1)} = 0$ (and $\neq 0$), thus it means that the effect of x_i is non-existent (and existent homogeneously in the two series of binary data). $a_i = 2$ means that the effect is existent and heterogeneous in the two.

In Table 4, the level indicates the number of variables included. As in the previous section, the dummy variable should be always included whether they are homogeneous (a_1= 1) or heterogeneous (a_1=2). Thus in Table 4, there are two models in level 1. In level 2, there are $2 \times 2 \times {}_3C_1$ =12 models, and in levels 3 and 4 the numbers of models are $2\times 2^2 \times {}_3C_2 = 24$ and $2\times 2^3 \times {}_3C_3 = 16$, respectively. Table 4 exhausts all possible models plausible in the probability model (3,1) with (3.2) as the least simple one.

In this combined hierarchy, the null models, corresponding to the null hypothesis of statistical independence, are (1 0 0 0), (1 0 0 1), (1 0 1 0), (1 1 0 0), (1 0 1 1), (1 1 0 1), (1 1 1 0) and (1 1 1 1).

The set of the null models itself forms a hierarchy and is isomorphic to the hierarchy in the previous section, and we call it the major hierarchy. To each one of the non-null models there are associated one or more models in the major hierarchy, which can be obtained by simplifying the model, and amomg them there is one which is the least simple one and is called the pivotal model. In other words, the combined hierarchy is partitioned as the sum of disjoint subsets, each one of which is represented by its pivotal model and forms again a hierarchy. The relations in the same subset are called the minor hierarchial relations and among them those involving the pivotal model the pivotal relations.

The hypothesis of independence justifies our nomenclatures and gives rise to the following rule. At first examine the relations in the major hierarchy and then proceed to the minor hierarchial relations, and among the latter we should begin with the pivotal relations and go to the rest of them.

Now let us return to Table 4. In the first part, all the models in the combined hierarchy down through (1 1 1 1) are listed to examine the goodness of fit χ^2 and also the decreases in χ^2 due to the inclusion of new variables. All of the models in the major hierarchy are rejected because of their high χ^2 values for testing the goodness of fit statistics.

Further let us proceed to the second part of Table 4, where all the models with (1 1 1 1) as the simplest one are listed and examined. All the pivotal relations are significant except for the one down to (1 1 1 2), which is also rejected. Among those with 19 degrees of freedom, the model (2 1 1 1) has the smallest χ^2 value and the largest log-likelihood value, and all the minor relations down from it are not significant. Thus we may conclude that the model (2 1 1 1) is the most appropriate one. The significant outcome in the test of goodness of fit for (1 1 1 2) may be judged as spurious because of two reasons. At first the decrease in the χ^2 value is not significant, and it attains the maximum χ^2 value among those with 19 degrees of freedom in this minor hierarchy.

The maximum likelihood estimates of the logit transforms for the model (2 1 1 1) are given by

$$\begin{cases} \lambda^{(0)} = \log(P^{(0)}/(1-P^{(0)})) = -2.531 + 0.171x_2 + 0.228x_3 - 0.513x_4 \\ \lambda^{(1)} = \log(P^{(1)}/(1-P^{(1)})) = -2.531 + 0.171x_2 + 0.228x_3 - 0.513x_4 \end{cases} \quad (3.3)$$

The outcomes in Table 4 are not discordant with those in the previous section. The distinction between the major and minor hierarchies is based on the hypothesis of independence. Without this distinction, we would be in a confusing situation, as the model (2 0 0 0) and many of its less simple ones, including (2 0 2 0), corresponding to the model judged as adequate in Table 2, appear to be equally adequate.

4. Discussions

The main purpose of this paper is to demonstrate the logical aspects of determining one or more adequate models, after performing all plausible analyses, which amount to quite large in number. We used the terminology "hierarchy", which needs some more discussions here.

Given a family of sets of linear models, such as (3.2), involving the same independent variables, where the regression coefficients on an independent variable are heterogeneous or homogeneous or non-existent, we have a natural partial order among them. In a pair of sets of models, one may be obtained from the other by deleting some part of the parameters and by assuming the homogeneity among the coefficients to some other set of variables, and in this case the one is said to be simpler than the other. The terminology "simpler" introduces a partial order. It is easy to verify that the family forms a lattice, including the maximum and the minimum in it. The use of the terminology "hierarchy" is based on the partial order with an associated test of decrease in the χ^2 values of the models in a pair.

In our present case, we have a compound hypothesis, which corresponds to a sub-lattice, and we called it the major hierarchy. For any element in the hierarchy, there corresponds uniquely an element in the major hierarchy which is the least simple one among those simpler than it. This mapping creates a partition of the lattice. Each of the set in this partition forms again a sub-lattice and the partial orders in these sub-lattices are minor hierarchial relations. Among the minor relations, those between an element of the major hierarchy and another in its minor hierarchy are called the pivotal relations.

The rule we are suggesting in this paper may be enunciated in the following: (a) It is the fundamental requirement that a model to be judged as adequate should be as simple as possible, in other words, the degrees of freedom should be as large as possible. (b) Examinations of the outcomes of the tests of goodness of fit are to be made in the order of (1) the major, (2) the pivotal and (3) the minor hierarchial relations. (c) Starting from the simplest model, significances in the tests of goodness of fit are to be examined successively towards the less simple ones in the order stated in (b), until we arrive at a model which exhibits no significance. (d) While examining the goodness of fit, attentions should be paid to the decreases in the χ^2 values due to going to the less simple models, although it plays the secondary role in the judgement.
(e) It is also to be kept in mind that in a series of tests, dependent as well as independent, with the same degrees of freedom, the maximum χ^2 tends to exhibit significance and the minimum non-significance.

The above stated rule still leaves a large amount of possibility of being unable to arrive at a conclusion and/or arriving at a multiplicity of possibly adequate models. Also this rule may have to be supplemented by some more requirements and may even be revised. As a matter of fact, this rule has occured with us not from logical reasonings, but from empirical observations of various data including the one in this paper.

In spite of these shortcomings, the present authors feel that the rule deserves to be spelled out here. The situation like the one stated at the end of the last section has occured with us not infrequently. It is also the belief of the authors that before we worry about establishing a complete and final rule, we should be more concerned with the quality and the limitations of the data such as in Table 1 and with the adequacy of the fundamental assumptions which generate the family of models concerned.

Further it is noted that no account has been taken of the correction for continuity in evaluating goodness of fit χ^2 statistics, in spite of the fact that all the n_i's are rather small for the data in Table 1. The smallness of the data is the main reason for the rather conservative attitude of the present authors. They decline to draw a decisive conclusion out of this analysis. Needless to say that the non-significance of the effect of C.E.E. grades as shown in (2.9) or (3.3) is meant to advocate neither abolishing nor continuing the college entrance examinations currently existing in our country. Nevertheless they feel that it still serves as a good demonstration of the method of analysis.

Finally the following issues are ignored or tacitly assumed in this paper. (cf. [4]) (1) Whether does the system of likelihood equations appearing in this paper have a solution or not ? (2) Is the solution, if exists, unique ? (3) Does the solution really maximize the likelihood function ? (4) Does the iterative method of Newton-Raphson really lead us to the solution ? (5) What is an efficient computor algorithm for computing all the statistics concerned ?

Appendix

$$\frac{\partial L}{\partial \beta_{au}} = t_{au} - \sum_i n_i x_{iu} P_{ai}$$

$$\frac{\partial^2 L}{\partial \beta_{au} \partial \beta_{av}} = - \sum_i n_i x_{iu} x_{iv} P_{ai} (1-P_{ai})$$

$$\frac{\partial^2 L}{\partial \beta_{au} \partial \beta_{bv}} = \sum_i n_i x_{iu} x_{iv} P_{ai} P_{bi}$$

$a,b = 1,2,3, \quad a \neq b; \quad u,v = 1,2,3,4$

$$\frac{\partial L_I}{\partial \beta_{au}} = t_{au} + t_{3u} - \sum_i n_i x_{iu} (P_{ai} + P_{3i})$$

$$\frac{\partial^2 L_I}{\partial \beta_{au} \partial \beta_{av}} = - \sum_i n_i x_{iu} x_{iv} (P_{ai} + P_{3i})(1 - P_{ai} - P_{3i})$$

$$\frac{\partial^2 L_I}{\partial \beta_{au} \partial \beta_{bv}} = 0$$

$a,b = 1,2, \quad a \neq b; \quad u,v = 1,2,3,4$

REFERENCES

[1] Bock, R. D. (1970). "Estimating multinomial response relations," Essays in Probability and Statistics, University of North Carolina Press, 111-132.

[2] Cox, D. R. (1966). "Some procedures connected with the logistic qualitative response curve," Research Papers in Statistics, John Wiley, 55-71.

[3] Cox, D. R. (1970). The Analysis of Binary Data. Methuen & Co., Ltd.

[4] Haberman, S. J. (1974). The Analysis of Frequency Data. The University of Chicago Press.

[5] Ito, K. & Kudô, A. (1974). "Some logical issues in interpreting multivariate data by means of regression analysis," To appear in the Proceedings of the 8-th International Biometric Conference.

[6] Kudô, K. & Ito, K. (1974). "Certain statistical methods in the analysis of the effect of inbreeding on mortality," To appear in the Proceedings of the 8-th International Biometric Conference.

[7] Mantel, N. (1966). "Models for complex contingency tables and polychotomous dosage response curves," Biometrics, 22, 83-95.

[8] Mantel, N. & Brown, C. (1973). "A logistic reanalysis of Ashford and Sowden's data on respiratory symptoms in British coal miners," Biometrics, 29, 649-665.

[9] Neyman, J. (1949). "Contributions to the theory of the χ^2 test," Proc. Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 239-273.

Table 1. Numbers of College Graduates According to Performances in Terms of Sex, High School Grades and College Entrance Examination Grades

Male

H. S. Grades		A			B			C		
C.E.E.Grades		A	B	C	A	B	C	A	B	C
(Y_1, Y_2)	(1, 1)	0	0	0	0	1	1	0	0	0
	(1, 0)	0	0	0	0	1	2	0	0	0
	(0, 1)	0	0	0	1	1	1	1	1	1
	(0, 0)	0	0	0	6	11	5	9	11	18
Total		0	0	0	7	14	9	9	12	19

Female

H. S. Grades		A			B			C		
C.E.E.Grades		A	B	C	A	B	C	A	B	C
(Y_1, Y_2)	(1, 1)	0	0	0	2	0	4	0	1	1
	(1, 0)	2	0	1	6	1	0	1	0	1
	(0, 1)	0	0	0	1	1	3	1	1	1
	(0, 0)	1	0	1	13	12	12	14	13	21
Total		3	0	2	22	14	19	16	15	24

Note: 1) Y_1= 1 if performances are excellent in general education,
0 otherwise,
Y_2= 1 if performances are excellent in major field,
0 otherwise.

2) The data were obtained from the registrar of a certain university for 1970 graduates from the faculty of arts and letters.

Table 2. Outcomes of Logistic Regression Analysis in Terms of Goodness of Fit χ^2 and Log Likelihood under No Assumption of Independence

Level	Model	χ^2	d.f.	Hierarchical relations, significance of decreases in χ^2 and Log-Likelihood
1	(1 0 0 0)	60.29*	39	X -135.18
	(1 0 0 1)	54.49*	36	N X -132.32
2 ⟹	(1 0 1 0)	26.59	36	S I . -124.24
	(1 1 0 0)	53.77*	36	N I I X -133.05
	(1 0 1 1)	22.86	33	S S N I . -121.86
3	(1 1 0 1)	53.33*	33	N N I N I X -130.35
	(1 1 1 0)	23.59	33	S I N S I I . -123.11
4	(1 1 1 1)	21.19	30	S S N S N S N . -120.77

Note: 1) In this table and in what follwos ., X, XX and XXX indicate that goodness of fit χ^2 are not significant, significant at 5%, 1% and 0.1% levels, respectively.

2) Each column in the lower triangle of the right portion of this table shows the relations of partial ordering, where S or N designates the decrease in χ^2 from the model corresponding to the top of the column is significant or not significant, and I designates that there is not a relation of partial ordering.

3) Numerical values for log likelihood are all minus the same constant.

4) ⟹ indicates that the so designated model is an adequate one.

Table 3. Outcomes of Logistic Regression Analysis in Terms of Goodness of Fit χ^2 and Log Likelihood under the Assumption of Independence

Level	Model	χ^2	d.f.	Hierarchical relations, significance of decreases in χ^2 and Log-Likelihood
1	(1 0 0 0)	99.53***	40	XXX -142.72
2	(1 0 0 1)	94.54***	38	N XXX -141.48
	(1 0 1 0)	66.29**	38	S I XX -131.90
	(1 1 0 0)	90.38***	38	S I I XXX -140.29
3	(1 0 1 1)	61.23**	36	S S N I XX -130.58
	(1 1 0 1)	85.77***	36	S S I N I XXX -139.08
	(1 1 1 0)	54.90*	36	S I S S I I X -130.59
4	(1 1 1 1)	49.80*	34	S S S S S S N X -129.24

Table 4a. Outcomes of Univariate Regression Analysis for Interpreting the Association between Y_1 and Y_2 in Terms of Goodness of Fit χ^2 and Log Likelihood

Model	χ^2	d.f.	Hierarchical relations, significance of decreases in χ^2 and Log Likelihood
(1 0 0 0)	49.92***	23	X -47.95
(2 0 0 0)	17.22	22	s . -41.80
(1 0 0 1)	43.26**	22	S i X -47.15
(2 0 0 1)	14.21	21	n n s . -40.97
(1 0 0 2)	43.16**	21	s i n i x -45.49
(2 0 0 2)	13.26	20	s n s n s . -40.41
(1 0 1 0)	45.13**	22	S i I i i i X -46.66
(2 0 1 0)	16.99	21	s n i i i i s . -41.69
(1 0 2 0)	29.22	21	s i i i i i s i . -43.63
(2 0 2 0)	14.26	20	s n i i i i s n s . -40.50
(1 1 0 0)	46.65**	22	N i I i i i I i i i X -47.68
(2 1 0 0)	17.16	21	s n i i i i i i i i s . -41.79
(1 2 0 0)	22.37	21	s i i i i i i i i i s i . -43.24
(2 2 0 0)	17.18	20	s n i i i i i i i i s n s . -41.78
(1 0 1 1)	39.08**	21	S i S i i i S i i i I i i i X -45.60
(2 0 1 1)	13.93	20	s n s n i i s n i i i i i i s . -40.74
(1 0 2 1)	23.53	20	s i s i i i s i s i i i i i s i . -42.66
(1 0 1 2)	36.43*	20	s i s i s i s i i i i i i i n i i x -44.20
(2 0 2 1)	11.29	19	s n s n i i s n s n i i i i s n s i . -39.65
(2 0 1 2)	13.05	19	s n s n s n s n i i i i i i s n i s i . -40.15
(1 0 2 2)	21.39	19	s i s i s i s i s i i i i i s i n s i i . -41.90
(2 0 2 2)	10.31	18	s n s n s n s n s n i i i i s n s s n n s . -39.26
(1 1 0 1)	40.32**	21	S i N i i i I i i i S i i i I i i i i i i i X -46.82
(2 1 0 1)	14.24	20	s n s n i i i i i i s n i i i i i i i i i i s . -40.96
(1 2 0 1)	18.74	20	s i s i i i i i i i s i n i i i i i i i i i i s i . -42.24
(1 1 0 2)	39.39**	20	s i n i n i i i i i s i i i i i i i i i i i i n i i x -45.21
(2 2 0 1)	14.14	19	s n s n i i i i i i s n s n i i i i i i i i i s n s i . -40.97
(2 1 0 2)	13.16	19	s n s n s n i i i i s n i i i i i i i i i i i s n i s i . -40.39
(1 2 0 2)	16.69	19	s i s i s i i i i i s i n i i i i i i i i i i s i n s i i . -41.07
(2 2 0 2)	13.07	18	s n s n s n i i i i s n s n i i i i i i i i i s n n s n n n . -40.35
(1 1 1 0)	42.66**	21	S i I i i i N i i i S i i i I i i i i i i i I i i i i i i i X -46.50
(2 1 1 0)	16.94	20	s n i i i i s n i i s n i i i i i i i i i i i i i i i i i i s . -41.68
(1 2 1 0)	21.24	20	s i i i i i s i i i s i n i i i i i i i i i i i i i i i i i s i . -42.88
(1 1 2 0)	28.71	20	s i i i i i s i n i s i i i i i i i i i i i i i i i i i i i s i i . -43.59
(2 2 1 0)	16.96	19	s n i i i i s n i i s n n n i i i i i i i i i i i i i i i i s n s i . -41.68
(2 1 2 0)	14.26	19	s n i i i i s n s n s n i i i i i i i i i i i i i i i i i i s n i s i . -40.50
(1 2 2 0)	17.23	19	s i i i i i s i s i s i n i i i i i i i i i i i i i i i i i s i s s i i . -41.45
(2 2 2 0)	14.29	18	s n i i i i s n s n s n s n i i i i i i i i i i i i i i i i s n s s n n n . -40.50
(1 1 1 1)	36.64*	20	S i S i i i S i i i S i i i N i i i i i i i N i i i i i i i S i i i i i i i X -45.40

Table 4b. Outcomes of Logistic Regression Analysis for Assessing the Association between Y_1 and Y_2 in Terms of Goodness of Fit χ^2 and Log Likelihood

Model	χ^2	d.f.	Hierarchical Relations and Significance of Decreases in χ^2
(1 1 1 1)	36.64*	20	X -45.40
(2 1 1 1)	13.94	19	s . -40.74
(1 2 1 1)	17.49	19	s i . -41.67
(1 1 2 1)	23.09	19	s i i . -42.61
(1 1 1 2)	33.85*	19	n i i i x -44.01
(2 2 1 1)	13.84	18	s n n i i . -40.74
(2 1 2 1)	11.28	18	s n i s i i . -39.66
(2 1 1 2)	12.95	18	s n i i s i i . -40.14
(1 2 2 1)	13.72	18	s i n s i i i i . -40.42
(1 2 1 2)	16.11	18	s i n i s i i i i . -40.52
(1 1 2 2)	20.91	18	s i i n s i i i i i . -41.84
(2 2 2 1)	11.23	17	s n s s i n n i n i i . -39.67
(2 2 1 2)	12.96	17	s n n i s n i n i n i i . -40.07
(2 1 2 2)	10.28	17	s n i s s i n n i i s i i . -39.26
(1 2 2 2)	12.18	17	s i n s s i i i n s s i i i . -39.68
(2 2 2 2)	10.24	16	s n n s s n n n n n s n n n n . -39.25

For the sake of convenience in printing XX and XXX are printed as X. Capital letters are used for the major relations, and small letters for the others.