

Normal Distribution Type Symmetry Model for Square Contingency Tables with Ordered Categories

Kouji Tahata, Kouji Yamamoto* and Sadao Tomizawa

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan

Abstract: For square contingency tables with ordered categories, this paper proposes a model that the cell probabilities have a similar structure of bivariate normal density function with equal marginal variances.

The proposed model is special cases of Agresti's linear diagonals-parameter symmetry model and Goodman's diamond model. Examples with real data are given and the simulation studies based on the bivariate normal distribution are also given.

Key Words: Diamond model, Linear diagonals-parameter symmetry, Normal distribution, Ordered category.

1. INTRODUCTION

Consider an $r \times r$ square contingency table with the same ordinal row and column classifications. Let p_{ij} denote the probability that an observation will fall in the i th row and j th column of the table ($i = 1, \dots, r; j = 1, \dots, r$). Agresti [1] proposed the linear diagonals-parameter symmetry (LDPS) model defined by

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

A special case of this model obtained by putting $\theta = 1$ is the usual symmetry model (see, for example, Bowker [2]; Bishop, Fienberg and Holland [3, p. 282]). This indicates that the probability that an observation will fall in the (i, j) th cell, $i < j$, is θ^{j-i} times higher than the probability that it falls in the (j, i) th cell.

Goodman [4] proposed the diagonals-parameter symmetry (DPS) model defined by

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

Note that the LDPS model is a special case of the DPS model obtained by putting $\{\xi_{j-i} = \xi^{j-i}\}$.

Consider the random variables X_1 and X_2 having a bivariate normal distribution with means $E(X_1) = \mu_1$ and

$E(X_2) = \mu_2$, variances $Var(X_1) = Var(X_2) = \sigma^2$, and correlation $Corr(X_1, X_2) = \rho$, which have the density function $f(x_1, x_2)$. Then we see

$$\frac{f(x_1, x_2)}{f(x_2, x_1)} = \exp\left(\frac{(x_2 - x_1)(\mu_2 - \mu_1)}{(1 - \rho)\sigma^2}\right).$$

Agresti [1] pointed out that the $f(x_1, x_2)/f(x_2, x_1)$ has the form $\theta^{x_2 - x_1}$ for some constant θ , and hence the LDPS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances (see also Tomizawa [5]; Yamamoto, Iwashita and Tomizawa [6]).

Goodman [7] proposed the diamond (DD) model defined by

$$p_{ij} = \delta_{i-j} \gamma_{i+j} \quad (i = 1, \dots, r; j = 1, \dots, r);$$

see also Tomizawa [8]. The DD model states that there is a structure of quasi-independence between the difference-diagonal classification (i.e., the difference between the row and column classifications) and the sum-diagonal classification (i.e., the sum between the row and column classifications).

Under the DD model we see

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

Therefore the DD model is a special case of the DPS model.

By the way, the density function $f(x_1, x_2)$ further can be expressed as

*Address correspondene to this author at the Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan;
E-mail: yamamoto_kouji@ymail.plala.or.jp

$$f(x_1, x_2) = c\alpha_1^{(x_1-x_2)^2} \alpha_2^{x_1-x_2} b_1^{(x_1+x_2)^2} b_2^{x_1+x_2}, \tag{1}$$

where

$$c = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left[-\frac{(\mu_1-\mu_2)^2}{4\sigma^2(1-\rho)} - \frac{(\mu_1+\mu_2)^2}{4\sigma^2(1+\rho)}\right],$$

$$a_1 = \exp\left(\frac{-1}{4\sigma^2(1-\rho)}\right),$$

$$a_2 = \exp\left(\frac{\mu_1-\mu_2}{2\sigma^2(1-\rho)}\right),$$

$$b_1 = \exp\left(\frac{-1}{4\sigma^2(1+\rho)}\right),$$

$$b_2 = \exp\left(\frac{\mu_1+\mu_2}{2\sigma^2(1+\rho)}\right).$$

Thus, the density function $f(x_1, x_2)$ has the form $g(x_1 - x_2)h(x_1 + x_2)$ and hence the DD model also may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances.

We are now interested in considering a model such that the cell probabilities $\{p_{ij}\}$ themselves have a similar structure of bivariate normal density function with the form of equation (1).

The purpose of this paper is to propose a new model which may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal variances.

The new model is different from the LDPS and DD models (Section 2).

2. NORMAL DISTRIBUTION TYPE SYMMETRY MODEL

For an $r \times r$ square table, consider a new model defined by

$$p_{ij} = \xi\alpha_1^{(i-j)^2} \alpha_2^{i-j} \beta_1^{(i+j)^2} \beta_2^{i+j} \quad (i = 1, \dots, r; j = 1, \dots, r). \tag{2}$$

It is easily seen that model (2) is a special case of the DD model because model (2) is expressed by using parameters dependent on the terms $(i - j)$ and $(i + j)$.

Since model (2) is similar to the form (1), we shall refer to model (2) as the normal distribution type symmetry (NDS) model. For the $r \times r$ table, let X and Y denote the row and column variables, respectively. Under the NDS model we see

$$\frac{P_{ij}}{P_{ji}} = (\alpha_2^2)^{i-j} \quad (i > j).$$

Therefore this model also has the structure of the LDPS model. So, the NDS model is also a special case of the LDPS model. Namely, under the NDS model we see that the prob-

ability that an observation will fall in the (i, j) th cell, $i > j$, is $(\alpha_2^2)^{i-j}$ times higher than the probability that it falls in the (j, i) th cell. Also, under the NDS model, $\alpha_2 > 1$ is equivalent to $P(X \leq i) < P(Y \leq i)$, $i = 1, \dots, r - 1$. Thus, the parameter α_2 in the NDS model would be useful for making inferences such as that X is stochastically greater than Y or vice versa.

Let Δ^* denote the $(2r - 1) \times (2r - 1)$ table of the diamond shape formed by rotating the $r \times r$ table based on the probabilities forty-five degrees. Table 1 shows the table Δ^* formed by rotating the original 4×4 table forty-five degrees. Generally, we could consider the $(2r - 1) \times (2r - 1)$ table of the diamond shape formed by rotating the original $r \times r$ table forty-five degrees so that the $2r - 1$ difference-diagonals in the original table now form the entries in the rows of the diamond, and the corresponding $2r - 1$ sum-diagonals in the original table form the entries in the columns of the diamond. Let $U = X - Y$ and $V = X + Y$. Also let S^* denote a set of cells of the diamond shape in the $(2r - 1) \times (2r - 1)$ table. Thus,

$$S^* = \{(u, v) \mid u = i - j, v = i + j \text{ for } i = 1, \dots, r; j = 1, \dots, r\}.$$

Table 1. Table Δ^* Formed by Rotating the Original 4 x 4 Table. Note that $p_{ij} = p^*_{uv}$ with $u = i - j$ and $v = i + j$

$U (= X - Y)$	$V (= X + Y)$						
	2	3	4	5	6	7	8
-3	*	*	*	p_{14}	*	*	*
-2	*	*	p_{13}	*	p_{24}	*	*
-1	*	p_{12}	*	p_{23}	*	p_{34}	*
0	p_{11}	*	p_{22}	*	p_{33}	*	p_{44}
1	*	p_{21}	*	p_{32}	*	p_{43}	*
2	*	*	p_{31}	*	p_{42}	*	*
3	*	*	*	p_{41}	*	*	*

Let p^*_{uv} denote the corresponding probability for row value u and column value v , $(u, v) \in S^*$, in the $(2r - 1) \times (2r - 1)$ table; i.e.,

$$P^*_{uv} = P_{\frac{u+v}{2}, \frac{v-u}{2}} \quad ((u, v) \in S^*).$$

The NDS model may be expressed as

$$P^*_{uv} = \xi\alpha_1^{u^2} \alpha_2^u \beta_1^{v^2} \beta_2^v \quad ((u, v) \in S^*).$$

Therefore under this model we see that there is a structure of quasi-independence between the difference-diagonal

classification (i.e., the difference between the row and column classifications, $U = X - Y$) and the sum-diagonal classification (i.e., the sum between the row and column classifications, $V = X + Y$).

The NDS model states that (1) when $V (= X + Y)$ is fixed as $V = v$, log-probability ($\log p_{uv}^*$) is a quadratic function of $U (= X - Y)$ and (2) when U is fixed as $U = u$, it is a quadratic function of V . In addition, (1) if $\alpha_1 > 1$ and $\alpha_2 > 1$, the probability p_{uv}^* for $u \geq 0$ increases as the value of U increases when the value of V is fixed, and (2) if $\beta_1 > 1$ and $\beta_2 > 1$, the probability p_{uv}^* increases as the value of V increases when the value of U is fixed.

The NDS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances. The fact is shown in Section 5 in terms of the simulation studies. For some details, see Section 5.

Also, a special case of the NDS model obtained by putting $\alpha_2 = 1$ may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal means and equal marginal variances.

3. GOODNESS-OF-FIT TEST

Let n_{ij} denote the observed frequency in the (i, j) th cell of the table ($i = 1, \dots, r; j = 1, \dots, r$). Assume that a multinomial distribution applies to the $r \times r$ table. The maximum likelihood estimates of expected frequencies under the NDS model can be easily obtained using an iterative procedure, for example, the general iterative procedure for log-linear models of Darroch and Ratcliff [9].

We shall consider the case where the sample size is large. The likelihood ratio statistic for testing the goodness-of-fit of a model symbolized by M is

$$G^2(M) = 2 \sum_{i=1}^r \sum_{j=1}^r n_{ij} \log \left(\frac{n_{ij}}{\hat{m}_{ij}} \right),$$

where \hat{m}_{ij} is the maximum likelihood estimate of expected frequency m_{ij} under model M . The number of degrees of freedom for the NDS model is $r^2 - 5$.

Consider two nested models, say M_1 and M_2 , such that model M_2 is a special case of model M_1 , so when M_2 holds, necessarily M_1 also holds. For example, M_2 is the NDS model and M_1 is the DD model. Let v_1 and v_2 denote the degrees of freedom for models M_1 and M_2 , respectively. Note that $v_1 < v_2$ and $G^2(M_1) \leq G^2(M_2)$. For testing that model M_2 holds assuming that model M_1 holds, we can use the likelihood ratio statistic $G^2(M_2 | M_1)$, where $G^2(M_2 | M_1) = G^2(M_2) - G^2(M_1)$. Under the null hypothe-

sis this test statistic has an asymptotic chi-square distribution with $v_2 - v_1$ degrees of freedom.

We note that (1) the NDS model implies the LDPS model, (2) the LDPS model implies the DPS model, (3) the NDS model also implies the DD model, and (4) the DD model implies the DPS model.

The readers may be interested in the case where the sample size is small. In such a case, we may be interested in considering the exact test. However, it may be difficult to consider the exact test for goodness-of-fit of the NDS model. In this paper, we focus on not testing goodness-of-fit for some models, but proposing a new model. Thus, although we do not discuss on the exact test in detail, for the exact test, for example, Aaberge [10] and Aaberge and Zhang [11].

4. AN EXAMPLE

Table 2 taken from Tomizawa, Miyamoto and Iwamoto [12] is the data of the decayed teeth of 363 women aged from 18 to 39 for the patients visiting a dental clinic in Sapporo City, Japan, from 2001 to 2005. Table 3 shows the table of the diamond shape formed by rotating the original table.

Table 2. Decayed Teeth Data of 363 Women Aged 18-39, for Patients Visiting a Dental Clinic in Sapporo City, Japan, from 2001 to 2005; from Tomizawa et al. [12]. (The Parenthesized Values are the Maximum likelihood Estimates of Expected Frequencies Under the NDS Model)

Left (Numbers of Decayed Teeth)	Right (Numbers of Decayed Teeth)			Total
	0-4 (1)	5-8 (2)	9+ (3)	
0-4 (1)	103 (103.41)	45 (45.46)	1 (2.72)	149
5-8 (2)	35 (33.53)	84 (84.40)	33 (28.90)	152
9+ (3)	3 (1.48)	17 (21.31)	42 (41.79)	62
Total	141	146	76	363

Table 3. Table of the Diamond Shape Formed by Rotating Table 2

Left Minus Right	Left Plus Right				
	2	3	4	5	6
-2	*	*	1	*	*
-1	*	45	*	33	*
0	103	*	84	*	42
1	*	35	*	17	*
2	*	*	3	*	*

From Table 4, we see that each model fits these data, and especially the NDS model is preferable to the other models in terms of likelihood ratio tests. Note that for the 3×3 table, the DD model is equivalent to the DPS model.

Table 4. Likelihood Ratio Statistic G^2 Values for Models Applied to Table 2

Applied Models	Degrees of Freedom	G^2
NDS	4	4.21
LDPS	2	4.18
DD	1	1.23
DPS	1	1.23

Under the NDS model, the values of maximum likelihood estimates of parameters are $\hat{\alpha}_1 = 0.393$, $\hat{\alpha}_2 = 0.859$, $\hat{\beta}_1 = 0.939$ and $\hat{\beta}_2 = 1.314$. Therefore, under the NDS model, the probability that the number of decayed teeth in the right side of the mouth of a patient is j and that in the left side is i ($< j$), is estimated to be $(1.356)^{j-i}$ ($= (1/\hat{\alpha}_2^2)^{j-i}$) times higher than the probability that the number of decayed teeth in the right side of the patient is i and that in the left side is j . Since $\hat{\alpha}_2 < 1$, the number of decayed teeth in the right side of the mouth of a patient tends to be more than that in the left side.

Also, the NDS model indicates that the total number of decayed teeth in the right and left sides of a patient is not associated with the difference between the number of decayed teeth in the right side and that in the left side of the patient.

5. SIMULATION STUDIES

As described in Section 2, the NDS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances. We shall now consider the simulation studies based on the bivariate normal distribution. Consider random variables Z_1 and Z_2 having a joint bivariate normal distribution with means $E(Z_1) = \mu_1$ and $E(Z_2) = \mu_2$, variances $Var(Z_1) = Var(Z_2) = \sigma^2$, and correlation $Corr(Z_1, Z_2) = \rho$. Suppose that there is an underlying bivariate normal distribution and suppose that a 4×4 table is formed using cutpoints for each variable at $\mu_1, \mu_1 \pm 0.6\sigma$. Then, in terms of simulation studies, Table 5 gives a 4×4 table of sample size 500, formed from an underlying bivariate normal distribution with the conditions $\sigma_1^2 = \sigma_2^2 (= \sigma^2)$, $\mu_2 = \mu_1 + 0.4$, and $\rho = 0.3$. The NDS model fits these data

well, yielding the likelihood ratio statistic $G^2 = 7.56$ with 11 degrees of freedom. Thus, the NDS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances.

As described in Section 1, the DD model may also be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances. Actually, the DD model fits the data in Table 5 well, yielding the likelihood ratio statistic $G^2 = 1.04$ with 4 degrees of freedom. However, according to the test based on the difference between the G^2 values for the NDS and DD models, (since the G^2 value is not significant at the 0.05 level under the assumption that the DD model holds), the NDS model is preferable to the DD model.

Table 5. The 4 x 4 table of sample size 500, formed by using cutpoints for each variable at $\mu_1, \mu_1 \pm 0.6\sigma$, from an underlying bivariate normal distribution with equal marginal variances and conditions $\mu_2 = \mu_1 + 0.4, \rho = 0.3$. (The parenthesized values are the maximum likelihood estimates of expected frequencies under the NDS model.)

Z_1	Z_2				Total
	(1)	(2)	(3)	(4)	
(1)	42 (37.00)	24 (30.03)	33 (31.62)	42 (43.21)	141
(2)	15 (17.92)	20 (18.46)	21 (24.66)	45 (42.75)	101
(3)	11 (11.26)	14 (14.72)	33 (24.95)	56 (54.87)	114
(4)	12 (9.19)	12 (15.23)	30 (32.75)	90 (91.38)	144
Total	80	70	117	233	500

By the way, as described in Section 1, the LDPS model may give a good fit when there is an underlying bivariate normal distribution with equal marginal variances. In fact, the LDPS model fits the data in Table 5 well, yielding the likelihood ratio statistic $G^2 = 2.00$ with 5 degrees of freedom. However, according to the test based on the difference between the G^2 values for the NDS and LDPS models, (since the G^2 value is not significant at the 0.05 level under the assumption that the LDPS model holds), the NDS model is preferable to the LDPS model. Thus, it may be preferable to assume the bivariate normal distribution on $\{p_{ij}\}$ rather than $\{p_{ij} / p_{ji}\}$.

In addition, consider the data in Table 6. Table 6 gives the simulated 4×4 square table of sample size 500 based on

the bivariate normal distribution with equal marginal variances and conditions $\mu_1 = \mu_2$, $\rho = 0.3$. The NDS model fits these data well, yielding the likelihood ratio statistic $G^2 = 7.35$ with 11 degrees of freedom. Under the NDS model applied to these data, the maximum likelihood estimates of α_1 , α_2 , β_1 and β_2 are $\hat{\alpha}_1 = 0.976$, $\hat{\alpha}_2 = 1.002$, $\hat{\beta}_1 = 1.081$ and $\hat{\beta}_2 = 0.465$. Hence, we see that $\hat{\alpha}_2$ is very close to 1. As described in Section 2, a special case of the NDS model obtained by putting $\alpha_2 = 1$ may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal means and equal marginal variances.

Table 6. The 4 x 4 Table of Sample Size 500, Formed by Using Cutpoints for Each Variable at $\mu_1, \mu_1 \pm 0.6\sigma$, from an Underlying Bivariate Normal Distribution with Equal Marginal Means and Equal Marginal Variances and $\rho = 0.3$. (The Parenthesized Values are the Maximum Likelihood Estimates of Expected Frequencies Under the NDS Model)

Z ₁	Z ₂				Total
	(1)	(2)	(3)	(4)	
(1)	57 (51.24)	32 (34.23)	27 (25.44)	17 (21.05)	133
(2)	27 (34.36)	30 (28.13)	30 (25.63)	31 (26.00)	118
(3)	22 (25.65)	25 (25.74)	26 (28.75)	32 (35.73)	105
(4)	23 (21.31)	31 (26.21)	36 (35.88)	54 (54.65)	144
Total	129	118	119	134	500

In addition, we perform many simulation studies under some conditions. For example, on conditions $\mu_2 = \mu_1 + 0.2$, $\sigma_1^2 = \sigma_2^2$, $\rho = 0.2$ and sample size 1000, the hypothesis that the NDS model holds is accepted 91 times per 100 times (at the 0.05 significant level), and also on conditions $\mu_2 = \mu_1 + 0.4$, $\sigma_1^2 = \sigma_2^2$, $\rho = 0.3$ and sample size 1000, that is accepted 94 times per 100 times. As a result of these many simulation studies, the NDS model tends to give a good fit when the marginal means are not so different and the correlation is not so large. However, it may be difficult to provide a limitation of difference between marginal means and value of correlation.

6. DISCUSSION

We have proposed the NDS model such that the cell probabilities $\{p_{ij}\}$ themselves have a similar structure of bivariate normal density function.

The NDS model is useful for making inferences such that the row variable X is stochastically greater than the column variable Y or vice versa, according to $\alpha_2 > 1$ or $\alpha_2 < 1$. In addition, under the NDS model, (i) if $\alpha_1 > 1$ and $\alpha_2 > 1$, the cell probability p_{ij} , $i \geq j$, tends to increase as the difference $i - j$ increases, when the sum $i + j$ is constant, and (ii) if $\beta_1 > 1$ and $\beta_2 > 1$, the cell probability p_{ij} tends to increase as the sum $i + j$ increases, when the difference $i - j$ is constant.

Finally, from Section 5 and many simulation studies, we can state that the NDS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances. In addition, when the marginal means equal, the NDS model with $\alpha_2 = 1$ would give a good fit to such data. As described in Section 5, it may be difficult to provide a clear guidance to use the NDS model. However, for analyzing a real contingency table data, if we may assume an underlying bivariate continuous distribution, it may be useful to apply the NDS model.

ACKNOWLEDGMENTS

The authors would like to express their thanks to referees for the helpful comments.

REFERENCES

- [1] A. Agresti, "A simple diagonals-parameter symmetry and quasi-symmetry model", *Statistics and Probability Letters*, vol. 1, pp. 313-316, 1983.
- [2] A. H. Bowker, "A test for symmetry in contingency tables", *Journal of the American Statistical Association*, vol. 43, pp. 572-574, 1948.
- [3] Y. M. M. Bishop, S. E. Fienberg, and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice*, Cambridge, Massachusetts: The MIT Press, 1975.
- [4] L. A. Goodman, "Multiplicative models for square contingency tables with ordered categories", *Biometrika*, vol. 66, pp. 413-418, 1979.
- [5] S. Tomizawa, "An extended linear diagonals-parameter symmetry model for square contingency tables with ordered categories", *Metron*, vol. 49, pp. 401-409, 1991.
- [6] H. Yamamoto, T. Iwashita, and S. Tomizawa, "Decomposition of symmetry into ordinal quasi-symmetry and marginal equipoment for multi-way tables", *Austrian Journal of Statistics*, vol. 36, pp. 291-306, 2007.
- [7] L. A. Goodman, "The analysis of cross-classified data having ordered and/or unordered categories: association models, correlation models, and asymmetry models for contingency tables with or without missing entries", *Annals of Statistics*, vol. 13, pp. 10-69, 1985.
- [8] S. Tomizawa, "Uniform association diamond model for square contingency tables with ordered categories", *Journal of the Japan Statistical Society*, vol. 24, pp. 83-88, 1994.
- [9] J. N. Darroch and D. Ratcliff, "Generalized iterative scaling for log-linear models", *Annals of Mathematical Statistics*, vol. 43, pp. 1470-1480, 1972.

- [10] R. Aaberge, "UMP unbiased tests for multiparameter testing problems with restricted alternatives", *Metrika*, vol. 50, pp. 179-193, 2000.
- [11] R. Aaberge, and L.-C. Zhang, "A class of exact UMP unbiased tests for conditional symmetry in small-sample square contingency tables", *Journal of Applied Statistics*, vol. 32, pp. 1-7, 2005.
- [12] S. Tomizawa, N. Miyamoto, and M. Iwamoto, "Linear column-parameter symmetry model for square contingency tables: application to decayed teeth data", *Biometrical Letters*, vol. 43, pp. 91-98, 2006.

Received: February 27, 2009

Revised: March 9, 2009

Accepted: March 31, 2009

© Tahata et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.