

# An Extended Bivariate T-Distribution Type Symmetry Model for Square Contingency Tables

Kiyotaka Iki\*, Masayuki Okada, Sadao Tomizawa

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Chiba, Japan  
Email: \*iki@is.noda.tus.ac.jp

**How to cite this paper:** Iki, K., Okada, M. and Tomizawa, S. (2018) An Extended Bivariate T-Distribution Type Symmetry Model for Square Contingency Tables. *Open Journal of Statistics*, 8, 249-257.  
<https://doi.org/10.4236/ojs.2018.82015>

**Received:** November 2, 2017

**Accepted:** April 1, 2018

**Published:** April 4, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The purpose of this paper is to propose a new model of asymmetry for square contingency tables with ordered categories. The new model may be appropriate for a square contingency table if it is reasonable to assume an underlying bivariate t-distribution with different marginal variances having any degrees of freedom. As the degrees of freedom becomes larger, the proposed model approaches the extended linear diagonals-parameter symmetry model, which may be appropriate for a square table if it is reasonable to assume an underlying bivariate normal distribution. The simulation study based on bivariate t-distribution is given. An example is given.

## Keywords

Bivariate T-Distribution, Square Contingency Table, Symmetry, Underlying Distribution

## 1. Introduction

Consider an  $R \times R$  square contingency table with the same row and column ordinal classifications. Let  $p_{ij}$  denote the probability that an observation will fall in the  $i$ th row and  $j$ th column of the table ( $i = 1, \dots, R; j = 1, \dots, R$ ). The symmetry (S) model is defined by

$$p_{ij} = p_{ji} \quad (i < j); \quad (1)$$

see Bowker [1] and Bishop *et al.* [2]. This model indicates a structure of symmetry of the probabilities with respect to the main diagonal of the table. Agresti [3] considered the linear diagonals-parameter symmetry (LDPS) model defined by

$$p_{ij} = \theta^{j-i} p_{ji} \quad (i < j). \quad (2)$$

This indicates that the probability that an observation will fall in the  $(i, j)$  th cell,  $i < j$ , is  $\theta^{j-i}$  times higher than the probability that it falls in the  $(j, i)$  th cell. A special case of the LDPS model obtained by putting  $\theta = 1$  is the S model. Furthermore, Tomizawa [4] proposed an extended linear diagonals-parameter symmetry (ELDPS) model defined by

$$p_{ij} = \theta_1^{j-i} \theta_2^{j^2-i^2} p_{ji} \quad (i < j). \tag{3}$$

This indicates that the probability that an observation will fall in the  $(i, j)$  th cell,  $i < j$ , is  $\theta_1^{j-i} \theta_2^{j^2-i^2}$  times higher than the probability that it falls in the  $(j, i)$  th cell.

Consider random variables  $X$  and  $Y$  having a joint bivariate normal distribution with means  $E(X) = \mu_1$ ,  $E(Y) = \mu_2$ , variances  $\text{Var}(X) = \sigma_1^2$ ,  $\text{Var}(Y) = \sigma_2^2$  and covariance  $\text{Cov}(X, Y) = \rho\sigma_1\sigma_2$ , where  $\rho$  is the correlation coefficient  $\text{Corr}(X, Y)$ . The joint bivariate normal density  $f(x, y)$  satisfies

$$\frac{f(x, y)}{f(y, x)} = \delta_1^{y-x} \delta_2^{y^2-x^2}, \tag{4}$$

where

$$\delta_1 = \exp \left[ \frac{1}{1-\rho^2} \left\{ \left( \frac{\mu_2}{\sigma_2^2} - \frac{\mu_1}{\sigma_1^2} \right) + \frac{\rho(\mu_2 - \mu_1)}{\sigma_1\sigma_2} \right\} \right],$$

$$\delta_2 = \exp \left[ \frac{1}{2(1-\rho^2)} \left( \frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) \right].$$

When  $\sigma_1^2 = \sigma_2^2 (= \sigma^2)$ ,  $f(x, y)$  satisfies

$$\frac{f(x, y)}{f(y, x)} = \delta^{y-x}, \tag{5}$$

where

$$\delta = \exp \left[ \frac{1}{1-\rho} \left( \frac{\mu_2 - \mu_1}{\sigma^2} \right) \right].$$

Agresti [3] [5] described the relationship between the LDPS model and the joint bivariate normal distribution as follows: 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. Also, Tomizawa [4] pointed out that the ELDPS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with different marginal variances.

Consider a bivariate t-distribution with  $m$  degree of freedom. The limit of this joint probability density function as  $m \rightarrow \infty$  is bivariate normal. Therefore, the LDPS model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with equal marginal variances such that  $m$  degree of freedom is very large. Also, the ELDPS model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with different marginal va-

riances such that  $m$  degree of freedom is very large.

Consider the  $R \times R$  square contingency table with ordered categories. For any fixed constant  $m (m > 2)$ , Iki *et al.* [6] proposed the t-distribution type symmetry (TS( $m$ )) model defined by

$$p_{ij}^{\frac{2}{m+2}} - p_{ji}^{\frac{2}{m+2}} = \eta_m (j-i) \quad (i < j). \quad (6)$$

A special case of this model obtained by putting  $\eta_m = 0$  is the S model. The TS( $m$ ) model indicates that the difference between the symmetric two probabilities raised to the power  $\left[ = -2/(m+2) \right]$  is proportional to the distance from the main diagonal of the  $R \times R$  table. The TS( $m$ ) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with equal marginal variances having  $m$  degrees of freedom (see Iki *et al.* [6]).

Now, we are interested in considering a new model, which is appropriate if it is reasonable to assume an underlying bivariate t-distribution with different marginal variances.

The purpose of present paper is to introduce an extended TS( $m$ ) model. Section 2 proposes an extended TS( $m$ ) model and describes the properties of the new model. Section 3 describes the relationships between the extended TS( $m$ ) model and t-distribution by the simulation study. Section 4 illustrates the use of our new model with father's and son's occupational structure data. Section 5 provides some concluding remarks.

## 2. Extended Bivariate T-Distribution Type Symmetry Model

Consider random variables  $X$  and  $Y$  having a joint bivariate t-distribution with  $m (m > 2)$  degrees of freedom, and means  $E(X) = \mu_1$ ,  $E(Y) = \mu_2$ , variances  $\text{Var}(X) = m\sigma_1^2/(m-2)$ ,  $\text{Var}(Y) = m\sigma_2^2/(m-2)$ , and correlation coefficient  $\text{Corr}(X, Y) = \rho$ . The probability density function  $h(x, y)$  is

$$h(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \left( 1 + \frac{Q(x, y)}{m} \right)^{\frac{m+2}{2}}, \quad (7)$$

where

$$Q(x, y) = \frac{1}{1-\rho^2} \left[ \left( \frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{x-\mu_1}{\sigma_1} \right) \left( \frac{y-\mu_2}{\sigma_2} \right) + \left( \frac{y-\mu_2}{\sigma_2} \right)^2 \right].$$

See, e.g., Muirhead [7]. Another form of the probability density function  $h(x, y)$  is expressed as

$$h(x, y) = c \left[ 1 + \frac{1}{m} (a_1x + b_1y + a_2x^2 + b_2y^2 + d(x, y)) \right]^{\frac{m+2}{2}}, \quad (8)$$

where

$$c = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

$$a_1 = \frac{2}{\sigma_1(1-\rho^2)} \left( \frac{\rho\mu_2}{\sigma_2} - \frac{\mu_1}{\sigma_1} \right), \quad b_1 = \frac{2}{\sigma_2(1-\rho^2)} \left( \frac{\rho\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2} \right),$$

$$a_2 = \frac{1}{\sigma_1^2(1-\rho^2)}, \quad b_2 = \frac{1}{\sigma_2^2(1-\rho^2)},$$

$$d(x, y) = \frac{1}{1-\rho^2} \left( -\frac{2\rho}{\sigma_1\sigma_2}xy + \frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} - \frac{2\rho\mu_1\mu_2}{\sigma_1\sigma_2} \right).$$

We note that  $d(x, y) = d(y, x)$ . Also, the probability density function  $h(x, y)$  satisfies

$$(h(x, y))^{\frac{2}{m+2}} - (h(y, x))^{\frac{2}{m+2}} = k_m(y^2 - x^2) + l_m(y - x) \quad (x < y), \tag{9}$$

where

$$k_m = \frac{\sigma_1^2 - \sigma_2^2}{m\sigma_1^2\sigma_2^2(1-\rho^2)} \left( 2\pi\sigma_1\sigma_2\sqrt{1-\rho^2} \right)^{\frac{2}{m+2}},$$

$$l_m = \frac{2\{\sigma_2^2\mu_1 - \sigma_1^2\mu_2 + \rho\sigma_1\sigma_2(\mu_1 - \mu_2)\}}{m\sigma_1^2\sigma_2^2(1-\rho^2)} \left( 2\pi\sigma_1\sigma_2\sqrt{1-\rho^2} \right)^{\frac{2}{m+2}}.$$

For continuous bivariate data, when we make the  $R \times R$  square contingency table formed using  $(R-1)$ 's cut points for each of row and column variables, we are interested in the structure of asymmetry of bivariate discrete probabilities  $\{p_{ij}\}$ . Consider the  $R \times R$  square contingency table with ordered categories. For any fixed constant  $m$ , we propose a model defined by

$$p_{ij} = \gamma \left[ 1 + \frac{1}{m}(\alpha_1 i + \beta_1 j + \alpha_2 i^2 + \beta_2 j^2 + \psi(i, j)) \right]^{\frac{m+2}{2}}, \tag{10}$$

where  $\psi(i, j) = \psi(j, i)$ . We shall refer to this model as an extended t-distribution type symmetry (ETS( $m$ )) model. The ETS( $m$ ) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with different marginal variances having  $m$  degrees of freedom. Under the ETS( $m$ ) model, setting  $\tau_{ij} = \alpha_1 i + \beta_1 j + \alpha_2 i^2 + \beta_2 j^2 + \psi(i, j)$ , we see that

$$\lim_{m \rightarrow \infty} \frac{p_{ij}}{p_{ji}} = \lim_{m \rightarrow \infty} \frac{\left( 1 + \frac{\tau_{ij}}{m} \right)^{-\frac{m+2}{2}}}{\left( 1 + \frac{\tau_{ji}}{m} \right)^{-\frac{m+2}{2}}} = \lim_{m \rightarrow \infty} \frac{\left\{ \left( 1 + \frac{\tau_{ij}}{m} \right)^{\frac{m}{\tau_{ij}}} \right\}^{-\frac{\tau_{ij}}{2} \left( 1 + \frac{2}{m} \right)}}{\left\{ \left( 1 + \frac{\tau_{ji}}{m} \right)^{\frac{m}{\tau_{ji}}} \right\}^{-\frac{\tau_{ji}}{2} \left( 1 + \frac{2}{m} \right)}} = \frac{\exp \left[ -\frac{\tau_{ij}}{2} \right]}{\exp \left[ -\frac{\tau_{ji}}{2} \right]} \tag{11}$$

$$= \exp \left[ \frac{1}{2}(\alpha_1 - \beta_1)(j - i) + \frac{1}{2}(\alpha_2 - \beta_2)(j^2 - i^2) \right] = \theta_1^{j-i} \theta_2^{j^2 - i^2},$$

where

$$\theta_1 = \exp \left[ \frac{1}{2}(\alpha_1 - \beta_1) \right],$$

$$\theta_2 = \exp \left[ \frac{1}{2}(\alpha_2 - \beta_2) \right].$$

Namely, the ETS( $m$ ) model approaches the ELDPs model as  $m$  becomes larg-

er, although the TS( $m$ ) model approaches the LDPS model as  $m$  becomes larger (see **Appendix 1**).

The ETS( $m$ ) model is also expressed as

$$p_{ij}^{\frac{2}{m+2}} - p_{ji}^{\frac{2}{m+2}} = \gamma_m (j^2 - i^2) + \eta_m (j - i) \quad (i < j), \quad (12)$$

where

$$\gamma_m = \frac{\gamma^{\frac{2}{m+2}} (\beta_2 - \alpha_2)}{m}, \quad \eta_m = \frac{\gamma^{\frac{2}{m+2}} (\beta_1 - \alpha_1)}{m}.$$

A special case of this model obtained by putting  $\gamma_m = 0$  is the TS( $m$ ) model.

The maximum likelihood estimates of expected frequencies under the ETS( $m$ ) model could be obtained using the Newton-Raphson method in the log-likelihood equation (see **Appendix 2**). For the ETS( $m$ ) model,  $\{p_{ij}\}$  are determined by  $R(R-1)/2$  of  $\{\psi(i, j), i < j\}$ ,  $R$  of  $\{\psi(i, i)\}$ , 1 of  $\gamma_m$  and 1 of  $\eta_m$ , thus a total of  $(R^2 + R + 4)/2$ . Therefore, the number of degrees of freedom (df) for the ETS( $m$ ) is  $R^2 - (R^2 + R + 4)/2 = (R^2 - R - 4)/2$ , which is one less than that for the TS( $m$ ) model, and equal to that for the ELDPS model.

### 3. Simulation Study

As described in Section 2, the ETS( $m$ ) model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate t-distribution with different marginal variances having  $m$  degrees of freedom. We shall consider the simulation study based on bivariate t-distribution. Consider random variables  $X$  and  $Y$  having a bivariate t-distribution with  $m(m > 2)$  degrees of freedom, and means  $E(X) = \mu_1$ ,  $E(Y) = \mu_2 (= \mu_1 + 0.2)$ , variances  $\text{Var}(X) = m\sigma_1^2/(m-2)$ ,  $\text{Var}(Y) = m\sigma_2^2/(m-2)$  and correlation coefficient  $\text{Corr}(X, Y) = \rho (= 0.2)$ . (Note that it is possible to take various values of  $\rho$ .) Such random numbers are obtained by using normal random number and chi-square random number with  $m$  degree of freedom. Suppose that there is an underlying bivariate t-distribution with some conditions, namely,  $\sigma_2^2/\sigma_1^2 = 1.2, 1.5$  and  $1.7$ , and a  $4 \times 4$  table of sample size 1000 is formed using cut points for each variable at  $\mu_1, \mu_1 \pm 0.6\sigma_1$ .

Then, we shall count the frequencies of acceptance (at the 0.05 significance level) based on the likelihood ratio chi-squared statistic for testing the hypothesis that the ETS( $m$ ) model or the TS( $m$ ) model with the corresponding  $m$  degrees of freedom of underlying t-distribution holds per 10000 times for  $4 \times 4$  tables on each conditions. From **Table 1**, we see that the ETS( $m$ ) model gives good fit, however, the TS( $m$ ) model gives poor fit. Thus, from the result of simulation for comparison the ETS( $m$ ) and TS( $m$ ) models, we obtain that if it is reasonable to assume the underlying bivariate t-distribution with different marginal variances and the ratio of different marginal variances (*i.e.*,  $\sigma_2^2/\sigma_1^2$ ) being large, the corresponding ETS( $m$ ) model rather than the TS( $m$ ) model would fit the data well without depending the value of degrees of freedom  $m$ .

**Table 1.** The frequencies of acceptance (at the 0.05 significance level) based on the likelihood ratio chi-squared statistic for testing the hypothesis that the ETS( $m$ ) model or the TS( $m$ ) model with the corresponding degrees of freedom of underlying t-distribution holds per 10,000 times for  $4 \times 4$  tables on some ratios of variances  $\sigma_2^2/\sigma_1^2$  and  $m$  degrees of freedom.

(a) For the ETS( $m$ ) model		
$\sigma_2^2/\sigma_1^2$	$m$	Frequencies
1.2	5	9323
1.2	20	9442
1.2	50	9418
1.2	100	9459
1.5	5	9374
1.5	20	9413
1.5	50	9451
1.5	100	9412
1.7	5	9407
1.7	20	9430
1.7	50	9424
1.7	100	9421
(b) For the TS( $m$ ) model		
$\sigma_2^2/\sigma_1^2$	$m$	Frequencies
1.2	5	1486
1.2	20	1725
1.2	50	1558
1.2	100	1555
1.5	5	0
1.5	20	0
1.5	50	0
1.5	100	0
1.7	5	0
1.7	20	0
1.7	50	0
1.7	100	0

#### 4. Example

The data in **Table 2** is taken from Tominaga [8]. These data describe the cross-classification of father's and son's occupational status categories in Japan, which were examined in 1955. From **Table 3**, we see that the TS( $m$ ) ( $m = 5, 20, 50, 100$ ) models fit these data poorly, however, the ETS( $m$ ) ( $m = 5, 20, 50, 100$ ) and the ELDPS models fit these data well. We obtain the similar result without depending the value of degree of freedom  $m$ , although the detail is omitted. We also see that the values of the likelihood ratio chi-squared statistic  $G^2$  for the ETS( $m$ ) model approach the value of  $G^2$  for the ELDPS model as

**Table 2.** The cross-classification of father's and son's occupational status categories in Japan, which were examined in 1955 from Tominaga [8]. (The parenthesized values are the maximum likelihood estimates of expected frequencies under the ETS(5) model.)

Father's status	Son's status				Total
	(1)	(2)	(3)	(4)	
(1)	80 (80.00)	72 (72.50)	37 (33.76)	19 (24.96)	208
(2)	44 (43.42)	155 (155.00)	61 (54.65)	31 (30.35)	291
(3)	26 (29.38)	73 (78.72)	218 (218.00)	45 (43.14)	362
(4)	69 (64.45)	156 (156.41)	166 (167.26)	614 (614.00)	1005
Total	219	456	482	709	1866

Note: (1) is upper non-manual, (2) lower non-manual, (3) manual, and (4) agriculture.

**Table 3.** Likelihood ratio chi-squared values  $G^2$  for models applied to Table 2.

Applied models	df	$G^2$
TS (5)	5	80.64*
TS (20)	5	91.58*
TS (50)	5	94.76*
TS (100)	5	95.92*
ETS (5)	4	3.82
ETS (20)	4	2.68
ETS (50)	4	2.43
ETS (100)	4	2.35
ELDPS	4	2.27

\*means significant at the 0.05 level.

the value of  $m$  increases. Under the ETS(5) model, the maximum likelihood estimates of  $\gamma_5$  and  $\eta_5$  are  $\hat{\gamma}_5 = 0.335$  and  $\hat{\eta}_5 = -1.405$ , respectively. Therefore, the difference between the probability raised to the power  $-\frac{2}{m+2}$  [ $= -0.286$ ] that the occupational status category of the father in pair is  $i$  and that of his son is  $j$  ( $i < j$ ) and the probability raised to the power that the occupational status category of the father in pair is  $j$  and that of his son is  $i$ , is  $\hat{\gamma}_5(j^2 - i^2) + \hat{\eta}_5(j - i)$  [ $= 0.335 \times (j^2 - i^2) - 1.405 \times (j - i)$ ]. For example, the difference between the probability raised to the power that the occupational status category of the father in pair is (1) "upper non-manual" and that of his son is (2) "lower non-manual" and the probability raised to the power that the occupational status category of the father in pair is (2) "lower non-manual" and that of his son

is (1) “upper non-manual”, is estimated to be  $-0.400 [= 0.335 \times 3 - 1.405 \times 1]$ . Namely, the probability that the occupational status category of the father in pair is (1) “upper non-manual” and that of his son is (2) “lower non-manual” is estimated to be greater than the probability that the occupational status category of the father in pair is (2) “lower non-manual” and that of his son is (1) “upper non-manual”. Also, the probability that the occupational status category of the father in pair is (1) and that of his son is (3) is estimated to be greater than the probability that the occupational status category of the father in pair is (3) and that of his son is (1). However, for the other symmetric two probabilities with respect to the main diagonal, the probability that the occupational status category of the father in pair is  $j$  and that of his son is  $i$  ( $i < j$ ) is estimated to be greater than the probability that the occupational status category of the father in pair is  $i$  and that of his son is  $j$ .

## 5. Concluding Remarks

From the result of simulation studies, the  $ETS(m)$  model would be appropriate for a square contingency table if it is reasonable to assume an underlying bivariate t-distribution with different marginal variances, although, the  $TS(m)$  model may be appropriate for a square contingency table if it is reasonable to assume an underlying bivariate t-distribution with equal marginal variances having  $m$  degrees of freedom.

## Acknowledgements

The authors would like to thank referees for helpful comments.

## References

- [1] Bowker, A.H. (1948) A Test for Symmetry in Contingency Tables. *Journal of the American Statistical Association*, **43**, 572-574.  
<https://doi.org/10.1080/01621459.1948.10483284>
- [2] Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. The MIT Press, Cambridge.
- [3] Agresti, A. (1983) A Simple Diagonals-Parameter Symmetry and Quasi-Symmetry Model. *Statistics and Probability Letters*, **1**, 313-316.  
[https://doi.org/10.1016/0167-7152\(83\)90051-2](https://doi.org/10.1016/0167-7152(83)90051-2)
- [4] Tomizawa, S. (1991) An Extended Linear Diagonals-Parameter Symmetry Model for Square Contingency Tables with Ordered Categories. *Metron: International Journal of Statistics*, **49**, 401-409.
- [5] Agresti, A. (1984) *Analysis of Ordinal Categorical Data*. John Wiley & Sons, New York.
- [6] Iki, K., Ishihara, T. and Tomizawa, S. (2013) Bivariate t-Distribution Type Symmetry Model for Square Contingency Tables with Ordered Categories. *Model Assisted Statistics and Applications: An International Journal*, **8**, 315-319.
- [7] Muirhead, R.J. (2005) *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, New Jersey.
- [8] Tominaga, K. (1979) *Nippon no Kaisou Kouzou (Japanese Hierarchical Structure)*. University of Tokyo Press, Tokyo. (In Japanese)



## Appendix 1

The TS( $m$ ) model is also expressed as

$$p_{ij} = \gamma \left[ 1 + \frac{1}{m} (\alpha i + \beta j + \phi(i, j)) \right]^{\frac{m+2}{2}}, \quad (\text{A.1})$$

where  $\phi(i, j) = \phi(j, i)$ . Under the TS( $m$ ) model, we see that

$$\lim_{m \rightarrow \infty} \frac{p_{ij}}{p_{ji}} = \theta^{j-i}, \quad (\text{A.2})$$

where

$$\theta = \exp \left[ \frac{1}{2} (\alpha - \beta) \right].$$

Namely, the TS( $m$ ) model approaches the LDPS model as  $m$  becomes larger.

## Appendix 2

Let  $n_{ij}$  denote the observed frequency in the  $i$ th row and  $j$ th column of the table ( $i = 1, \dots, R; j = 1, \dots, R$ ), with  $n = \sum \sum n_{ij}$ . Assume that a multinomial distribution applies to the  $R \times R$  table. We consider the maximum likelihood estimates of expected frequencies  $\{m_{ij}\}$  under the ETS( $m$ ) model. We must maximize the Lgrangian

$$L = \sum_{i=1}^R \sum_{j=1}^R n_{ij} \log p_{ij} - \lambda \left( \sum_{i=1}^R \sum_{j=1}^R p_{ij} - 1 \right) - \sum_{i < j} \phi_{ij} \left\{ p_{ij}^{\frac{2}{m+2}} - p_{ji}^{\frac{2}{m+2}} - \gamma_m (j^2 - i^2) - \eta_m (j - i) \right\}, \quad (\text{A.3})$$

with respect to  $\{p_{ij}\}, \lambda, \{\phi_{ij}\}, \gamma_m$  and  $\eta_m$ . Setting the partial derivations of  $L$  equal to zero, we obtain the equations:

$$\frac{\partial L}{\partial p_{kl}} = \frac{n_{kl}}{p_{kl}} - n + \frac{2}{m+2} \phi_{kl} p_{kl}^{\frac{m+4}{m+2}} (= 0) \quad (k < l), \quad (\text{A.4})$$

$$\frac{\partial L}{\partial p_{kl}} = \frac{n_{kl}}{p_{kl}} - n - \frac{2}{m+2} \phi_{lk} p_{kl}^{\frac{m+4}{m+2}} (= 0) \quad (k > l), \quad (\text{A.5})$$

$$\frac{\partial L}{\partial p_{kk}} = \frac{n_{kk}}{p_{kk}} - n (= 0) \quad (k = 1, \dots, R), \quad (\text{A.6})$$

$$p_{kl}^{\frac{2}{m+2}} - p_{lk}^{\frac{2}{m+2}} = \gamma_m (l^2 - k^2) + \eta_m (l - k) \quad (k < l), \quad (\text{A.7})$$

$$\sum_{i < j} \phi_{ij} (j^2 - i^2) = 0, \quad (\text{A.8})$$

$$\sum_{i < j} \phi_{ij} (j - i) = 0. \quad (\text{A.9})$$

Using the Newton-Raphson method, we can solve Equations (A.4) to (A.9) with respect to  $\{p_{ij}\}, \gamma_m$  and  $\eta_m$ . Noting that  $\{m_{ij} = np_{ij}\}$ , we obtain the maximum likelihood estimates of  $\{m_{ij}\}$  and the parameters  $\gamma_m$  and  $\eta_m$  under the ETS( $m$ ) model.