

Generalized Exponential Symmetry Model and Orthogonal Decomposition of Symmetry for Square Tables

Hiroyuki Kurakami*, 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 the analysis of square contingency tables with ordered categories, some models that the log odds for two symmetric cell probabilities is a linear function of the row and column values have been considered. This paper proposes a generalization of these models. This paper also proposes the model that the weighted sum of the probability that an observation will fall in one of the cells in upper right triangle of the table is equal to the weighted sum of the probability that it falls in one of the cells in lower left triangle of the table. In addition, this paper gives the theorem that the symmetry model is equivalent to both the proposed models holding simultaneously. Moreover, this paper shows that the likelihood ratio statistic for testing goodness-of-fit of the symmetry model is asymptotically equivalent to the sum of those for testing the proposed models. Examples are given.

Keywords: Generalized exponential symmetry, generalized weighted global symmetry, likelihood ratio statistic, orthogonality, square contingency table, symmetry.

1. INTRODUCTION

Consider an $r \times r$ square contingency table with the same 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 ($i = 1, \dots, r; j = 1, \dots, r$). The symmetry (S) model (Bowker [1]) is defined by

$$p_{ij} = \psi_{ij} \quad (i = 1, \dots, r; j = 1, \dots, r),$$

where $\psi_{ij} = \psi_{ji}$ and $\{\psi_{ij}\}$ are unspecified (also see Bishop, Fienberg and Holland [2], p. 282). This model is also expressed as

$$p_{ij} = p_{ji} \quad (i \neq j).$$

This model states that the probability that an observation will fall in the (i, j) th cell, $i \neq j$, is equal to the probability that it falls in the symmetric (j, i) th cell.

Some extensions of the S model have been considered. The conditional symmetry (CS) model (McCullagh [3]) is defined by

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

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

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

Let X and Y denote the row and column variables, respectively. The CS model indicates the symmetry of the conditional probabilities such that $P(X = i, Y = j | X < Y)$ equals $P(X = j, Y = i | X > Y)$ ($i < j$). A special case of this model obtained by putting $\theta = 1$ is the S model.

The linear diagonals-parameter symmetry (LDPS) model (Agresti [4]) is defined by

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

where $\psi_{ij} = \psi_{ji}$ and $\{\psi_{ij}\}$ are unspecified. This model 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. A special case of this model obtained by putting $\theta = 1$ is the S model.

The linear column-parameter symmetry (LCPS) model (Tomizawa, Miyamoto and Iwamoto [5]) is defined by

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

where $\psi_{ij} = \psi_{ji}$ and $\{\psi_{ij}\}$ are unspecified. This model indicates that the log odds, $\log(p_{ij} / p_{ji})$ is a linear function of the column value j . A special case of this model obtained by putting $\theta = 1$ is the S model.

*Address correspondence to this author at the Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan; Tel: +81 (0) 4-7124-1501; Fax: +81 (0) 4-7123-9764; E-mail: h-kurakami@mti.biglobe.ne.jp

The CS, LDPS and LCPS models have the structure that the log odds, $\log(p_{ij} / p_{ji})$ is a linear function of the row and column values. We are now interested in a generalization of these models.

For square contingency tables, some statisticians gave decompositions of the S model into some models. The orthogonal decompositions, which mean that the test statistic for a model is (asymptotically) equivalent to the sum of those for decomposed models, have been proposed, see, e.g., Read [6], Tomizawa and Tahata [7] and Tahata, Yamamoto and Tomizawa [8]. We are also interested in the orthogonal decomposition of the S model using the generalized model.

The purpose of this paper is (i) to propose the generalized exponential symmetry (GES) model, which indicates that the log odds for two symmetric cell probabilities is an arbitrary function of the row and column values, (ii) to give the decomposition of the S model using the GES model, and (iii) to show that the test statistic for the S model is asymptotically equivalent to the sum of those for decomposed models. Section 2 proposes the GES model, and gives the decomposition of the S model. Section 3 shows the orthogonality of the decomposition with respect to the goodness-of-fit test statistic. Section 4 gives examples.

2. MODELS AND DECOMPOSITION

For an $r \times r$ square contingency table with ordered categories, we consider the following model:

$$p_{ij} = \begin{cases} \theta^{w_{ij}} \psi_{ij} & (i < j), \\ \psi_{ij} & (i \geq j), \end{cases} \quad (1)$$

where $\psi_{ij} = \psi_{ji}$, $\{\psi_{ij}\}$ are unspecified and $\{w_{ij}\}$ are the specified nonnegative values. This indicates that the probability that an observation will fall in the (i, j) th cell, $i < j$, is $\theta^{w_{ij}}$ times higher than the probability that it falls in the (j, i) th cell. A special case of this model obtained by putting $\theta = 1$ is the S model. When $\{w_{ij} = 1\}$, $\{w_{ij} = j - i\}$ and $\{w_{ij} = j - 1\}$, this model is equivalent to the CS, LDPS and LCPS models, respectively. So, we shall refer to model (1) as the generalized exponential symmetry (GES) model.

Next, consider the following model:

$$\sum_{i=1}^{r-1} \sum_{j=i+1}^r w_{ij} p_{ij} = \sum_{j=1}^{r-1} \sum_{i=j+1}^r w_{ij} p_{ij}, \quad (2)$$

where $w_{ij} = w_{ji}$, $i < j$, and $\{w_{ij}\}$ are the specified nonnegative values. This model indicates that the weighted sum of the probability that an observation will fall in one of the cells in upper right triangle of the table is equal to the weighted sum of the probability that it falls in one of the cells in lower left triangle of the table. We shall refer to model (2) as the generalized weighted global symmetry (GWGS) model. When $\{w_{ij} = 1\}$, this model is identical to Read's [6] global symmetry (GS) model, and when $\{w_{ij} = j - i\}$, $i < j$, this model indicates the marginal mean

equality (ME), i.e., $E(X) = E(Y)$. Also, when $\{w_{ij} = j - 1\}$, $i < j$, model (2) indicates

$$\sum_{i=1}^{r-1} \sum_{j=i+1}^r (j-1) p_{ij} = \sum_{j=1}^{r-1} \sum_{i=j+1}^r (i-1) p_{ij}. \quad (3)$$

We shall refer to model (3) as the weighted global symmetry (WGS) model.

Then, we can obtain the decomposition of the S model as follows:

Theorem 1 : The S model holds if and only if the GES and GWGS models hold.

Proof : If the S model holds, then the GES and GWGS models hold. Assuming that both the GES and GWGS models hold, then we shall show that the S model holds.

From the assumption, we see

$$\sum_{i=1}^{r-1} \sum_{j=i+1}^r w_{ij} \theta^{w_{ij}} \psi_{ij} = \sum_{i=1}^{r-1} \sum_{j=i+1}^r w_{ij} \psi_{ij}.$$

Therefore, we obtain $\theta = 1$. Namely, the S model holds. The proof is completed.

3. ORTHOGONALITY OF THE DECOMPOSITION

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 is applied to the $r \times r$ table. The maximum likelihood estimates (MLEs) of expected frequencies under the GES and GWGS models could be obtained by using the iterative procedures, for example, the general iterative procedure for log-linear model of Darroch and Ratcliff [9] or using the Newton-Raphson method to the log-likelihood equation. Each model can be tested for goodness-of-fit by, e.g., the likelihood ratio chi-square statistic with the corresponding degrees of freedom (df). The likelihood ratio statistic for testing goodness-of-fit of model M is given by

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

where \hat{m}_{ij} is the MLE of expected frequency m_{ij} under model M .

For the analysis of contingency tables, Lang and Agresti [10] and Lang [11] considered the simultaneous modeling of the joint distribution and of the marginal distribution. Aitchison [12] discussed the asymptotic separability, which is equivalent to the orthogonality in Read [6] and the independence in Darroch and Silvey [13], of test statistic for the goodness-of-fit of two models (also see Lang and Agresti [10]; Lang [11]; Tomizawa and Tahata [7]; Tahata and Tomizawa [14]).

For the $r \times r$ table, we shall consider the orthogonality (i.e., separability or independence) of test statistic for the

decomposition of the S model into the GES and GWGS models.

Theorem 2: For an $r \times r$ table, the following asymptotic equivalence holds:

$$G^2(S) \simeq G^2(GES) + G^2(GWGS),$$

Where “ \simeq ” means that the difference between the left-hand side and right-hand side converges in probability to zero.

Proof: For an $r \times r$ square table, the GES model may be expressed as

$$\log p_{ij} = \begin{cases} w_{ij}\beta_0 + \varphi_{ij} & (i < j), \\ \varphi_{ii} & (i = j), \\ -w_{ij}\beta_0 + \varphi_{ij} & (i > j), \end{cases} \quad (4)$$

where $w_{ij} = w_{ji}$ and $\varphi_{ij} = \varphi_{ji}$. Let

$$p = (p_{11}, \dots, p_{1r}, p_{21}, \dots, p_{2r}, \dots, p_{r1}, \dots, p_{rr})',$$

$$\beta = (\beta_0, \beta_1)',$$

where $'$ denotes the transpose, and where

$$\beta_1 = (\varphi_{11}, \varphi_{12}, \dots, \varphi_{1r}, \varphi_{22}, \dots, \varphi_{2r}, \dots, \varphi_{r-1,r-1}, \varphi_{r-1,r}, \varphi_{rr})',$$

is the $r(r+1)/2 \times 1$ vector. Then the GES model is expressed as

$$\log p = X\beta = (X_0, X_1)\beta,$$

where X is the $r^2 \times K$ matrix with $K = (r^2 + r + 2)/2$, and X_0 is the $r^2 \times 1$ vector with

$$X_0 = (x_{11}, \dots, x_{1r}, x_{21}, \dots, x_{2r}, \dots, x_{r1}, \dots, x_{rr})',$$

where

$$x_{ij} = \begin{cases} w_{ij} & (i < j), \\ 0 & (i = j), \\ -w_{ij} & (i > j), \end{cases}$$

with $w_{ij} = w_{ji}$, and X_1 is $r^2 \times r(r+1)/2$ matrix of 0 or 1 elements determined from (4). Note that the matrix X is full column rank which is K . In a similar manner to Haber [15], Lang and Agresti [10], we denote the linear space spanned by the columns of the matrix X by $S(X)$ with the dimension K . Note that $X_1 1_{r(r+1)/2} = 1_{r^2}$ where 1_t is the $t \times 1$ vector of 1 elements, thus $1_{r^2} \in S(X)$. Let U be an $r^2 \times d_1$, where $d_1 = r^2 - K = (r-2)(r+1)/2$, full column rank matrix such that the linear space spanned by the column of U , i.e., $S(U)$, is the orthogonal complement of the space $S(X)$. Thus, $U'X = O_{d_1, K}$, where O_{st} is the $s \times t$ zero matrix. Therefore the GES model is expressed as

$$h_1(p) = 0_{d_1},$$

where 0_{d_1} is the $d_1 \times 1$ zero vector, and

$$h_1(p) = U' \log p.$$

The GWGS model may be expressed as

$$h_2(p) = 0_{d_2},$$

where $d_2 = 1$,

$$h_2(p) = X_0' p.$$

Note that $X_0' U = 0_{d_1}'$. From Theorem 1, the S model may be expressed as

$$h_3(p) = 0_{d_3},$$

where $d_3 = d_1 + d_2 = r(r-1)/2$,

$$h_3 = (h_1', h_2')'.$$

Note that $h_s(p), s=1,2,3$, are the vectors of order $d_s \times 1$, and $d_s, s=1,2,3$, are the numbers of df for testing goodness-of-fit of the GES, GWGS and S models, respectively.

Let $H_s(p), s=1,2,3$, denote the $d_s \times r^2$ matrix of partial derivatives of $h_s(p)$ with respect to p , i.e., $H_s(p) = \partial h_s(p) / \partial p'$. Let $\Sigma(p) = \text{diag}(p) - pp'$, where $\text{diag}(p)$ denotes a diagonal matrix with i th component of p as i th diagonal component. We see that

$$H_1(p)p = U' 1_{r^2} = 0_{d_1},$$

$$H_1(p)\text{diag}(p) = U',$$

$$H_2(p) = X_0'.$$

Therefore we obtain

$$H_1(p)\Sigma(p)H_2(p)' = U'X_0 = 0_{d_1}.$$

Thus we obtain $\Delta_3 = \Delta_1 + \Delta_2$, where

$$\Delta_s = h_s(p)' [H_s(p)\Sigma(p)H_s(p)]^{-1} h_s(p).$$

From the asymptotic equivalence of the Wald statistic and the likelihood ratio statistic (Rao [16], Sec. 6e. 3; Darroch and Silvey [13]; Aitchison [12]), we obtain Theorem 2. The proof is completed.

Note that for the $r \times r$ table, Read [6] shows $G^2(S) = G^2(CS) + G^2(GS)$, and Tahata *et al.* [8] shows $G^2(S) \simeq G^2(LDPS) + G^2(ME)$. Theorem 2 is a generalization of these orthogonal decompositions.

A quick method for choosing the best-fitting model among different models is to use Akaike's [17] information criterion (AIC), which is defined as

Table 1. Decayed Teeth Data of 363 Women with Lower and Upper Teeth in Sapporo, Japan; from Tomizawa et al. [5]. The Parenthesized Values are the Maximum Likelihood Estimates of Expected Frequencies Under the LCPS Model

Lower	Upper			Total
	(1)	(2)	(3)	
(1)	97 (97.00)	62 (62.79)	15 (15.54)	174
(2)	20 (19.21)	63 (63.00)	75 (74.06)	158
(3)	2 (1.46)	6 (6.94)	23 (23.00)	31
Total	119	131	113	363

Table 2. Occupational Status for Father-Son Pairs in Melbourne, Australia; from Miller [19]. The Parenthesized Values are the Maximum Likelihood Estimates of Expected Frequencies Under the WGS Model

Father's Status	Son's Status				Total
	(1)	(2)	(3)	(4)	
(1)	16 (16.00)	11 (10.24)	11 (9.59)	7 (5.73)	45
(2)	5 (5.40)	10 (10.00)	3 (2.61)	4 (3.28)	22
(3)	2 (2.35)	2 (2.35)	10 (10.00)	6 (4.91)	20
(4)	0 (0.00)	9 (11.55)	7 (8.99)	18 (18.00)	34
Total	23	32	31	35	121

$AIC = -2(\text{maximum log likelihood}) + 2(\text{number of parameters}),$

for each model. For details see Konishi and Kitagawa [18]. This criterion gives the best-fitting model as the one with minimum AIC. Since only the difference between AICs is required when two models are compared, it is possible to ignore a common constant of AIC and we may use a modified AIC defined as

$$AIC^+ = G^2 - 2(\text{number of df}).$$

Thus, for the data, the model with the minimum AIC^+ (i.e., the minimum AIC) is the best-fitting model. This criterion will be used in the next section.

4. EXAMPLES

Example 1: Consider the data in Table 1, taken from Tomizawa et al. [5]. These are decayed teeth data of 363 women with lower and upper teeth in Sapporo, Japan. Each category means (1): 0-4 decayed teeth, (2): 5-8 decayed teeth and (3): above 9 decayed teeth. Table 4 gives the likelihood

ratio chi-square statistic for each model. From Table 4a, we see that the LCPS model fits the data in Table 1 very well, and the others fit poorly. In addition, the LCPS model has a minimum AIC^+ value. Therefore when we use AIC criterion, the LCPS model is the best-fitting model among the models. From Theorem 1, we can see that the poor fit of the S model is caused by the influence of the poor fit of the WGS model rather than the LCPS model. Under the LCPS model, the MLE of $\log \theta$ with asymptotic standard error given in parenthesis is $\log \hat{\theta} = 1.184$ (a.s.e. = 0.150) and the MLE of θ is $\hat{\theta} = 3.268$. Since $\hat{\theta} > 1$, we can estimate that $P(X < Y) > P(X > Y)$, that is, the number of a woman's upper decayed teeth tends to be more than that of her lower decayed teeth.

Example 2: Consider the data in Table 2, taken directly from Miller [19]. These are occupational mobility data for father-son pairs in Melbourne, Australia. Each category means (1): Employer and self-employed, (2): White collar, (3): Skilled and (4): Semi-skilled. From Table 4b, we see

Table 3. Occupational Status for Father-Son Pairs in Norway; from Miller [19]. The Parenthesized Values are the Maximum Likelihood Estimates of Expected Frequencies Under the LDPS Model

Father's Status	Son's Status				Total
	(1)	(2)	(3)	(4)	
(1)	95 (95.00)	41 (36.81)	1 (3.24)	10 (8.95)	147
(2)	66 (70.19)	160 (160.00)	14 (17.20)	16 (18.12)	256
(3)	14 (11.76)	36 (32.80)	20 (20.00)	19 (14.45)	89
(4)	61 (62.05)	68 (65.88)	23 (27.55)	97 (97.00)	249
Total	236	305	58	142	741

that the WGS model fits the data in Table 2 well, and the others fit poorly. Also, the WGS model has a minimum AIC^+ value. Thus, when we use AIC criterion, the WGS model is the best-fitting model among the models. From Theorem 1, we can see that the poor fit of the S model is caused by the influence of the poor fit of the LCPS model rather than the WGS model. Since the WGS model fits well, we can estimate that the weighted sum of the probability that a father's status is i and his son's status is $j (> i)$ is equal to the weighted sum of the probability that the father's status is j and his son's status is i .

Example 3: Consider the data in Table 3, taken directly from Miller [19]. These are occupational mobility data for father-son pairs in Norway. Each category means (1): Non-manual, (2): Working-class, (3): Primary workers and (4): Primary independent. From Table 4c, we see that the LDPS model fits the data in Table 3 well, and the others fit poorly. In addition, the LDPS model has a minimum AIC^+ value. Therefore when we use AIC criterion, the LDPS model is the best-fitting model among the models. From Theorem 1, the poor fit of the S model is caused by the influence of the poor

fit of the ME model rather than the LDPS model. Under the LDPS model, the MLE of $\log \theta$ is $\log \hat{\theta} = -0.646$ (a.s.e. = 0.075) and the MLE of θ is $\hat{\theta} = 0.524$. Since $\hat{\theta} < 1$, we can estimate that $P(X < Y) < P(X > Y)$, namely, a father's occupational status category tends to be greater than his son's status category.

5. CONCLUDING REMARKS

In this paper, we have proposed a generalized model including the S, CS, LDPS and LCPS models and also introduced a generalized model including the GS and ME models. In addition, we have given the decomposition of the S model and shown the orthogonality of test statistic.

We point out from Theorem 2 that for instance, the likelihood ratio statistic for testing goodness-of-fit of the S model assuming that the GES model holds true is $G^2(S) - G^2(GES)$ and this is asymptotically equivalent to the likelihood ratio statistic for testing goodness-of-fit of the GWGS model, i.e., $G^2(GWGS)$.

Table 4. Likelihood Ratio Values G^2 and AIC^+ Values for Each Model Applied to the Data in Tables 1, 2, and 3. The Parenthesized Values are the Number of Parameters for Each Model

(a) For Table 1

Models	df	G^2	p -value	AIC^+
S	3 (5)	103.33	< 0.001	97.33
CS	2 (6)	9.40	0.009	5.40
GS	1 (7)	93.93	< 0.001	91.93
LDPS	2 (6)	11.05	0.004	7.05
ME	1 (7)	81.20	< 0.001	79.20
LCPS	2 (6)	0.39	0.823	- 3.61
WGS	1 (7)	102.16	< 0.001	100.16

Table 4. contd...

(b) For Table 2

Models	df	G^2	p -value	AIC ⁺
S	6 (9)	21.12	0.002	9.12
CS	5 (10)	16.76	0.005	6.76
GS	1 (14)	4.36	0.037	2.36
LDPS	5 (10)	14.86	0.011	4.86
ME	1 (14)	6.52	0.011	4.52
LCPS	5 (10)	18.92	0.002	8.92
WGS	1 (14)	2.17	0.141	0.17

(c) For Table 3

Models	df	G^2	p -value	AIC ⁺
S	6 (9)	105.09	< 0.001	93.09
CS	5 (10)	26.70	< 0.001	16.70
GS	1 (14)	78.40	< 0.001	76.40
LDPS	5 (10)	6.74	0.241	- 3.26
ME	1 (14)	96.99	< 0.001	94.99
LCPS	5 (10)	19.61	0.001	9.61
WGS	1 (14)	84.73	< 0.001	82.73

ACKNOWLEDGMENT

The authors would like to thank two anonymous reviewers for their helpful comments.

REFERENCES

- [1] A. H. Bowker, "A test for symmetry in contingency tables", *Journal of the American Statistical Association*, vol. 43, pp. 572-574, 1948.
- [2] Y. M. M. Bishop, S. E. Fienberg, and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice*, Cambridge: The MIT Press, 1975.
- [3] P. McCullagh, "A class of parametric models for the analysis of square contingency tables with ordered categories," *Biometrika*, vol. 65, pp. 413-418, 1978.
- [4] A. Agresti, "A simple diagonals-parameter symmetry and quasi-symmetry model," *Statistics and Probability Letters*, vol. 1, pp. 313-316, 1983.
- [5] 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.
- [6] C. B. Read, "Partitioning chi-square in contingency tables: A teaching approach," *Communications in Statistics-Theory and Methods*, vol. 6, pp. 553-562, 1977.
- [7] S. Tomizawa and K. Tahata, "The analysis of symmetry and asymmetry: Orthogonality of decomposition of symmetry into quasi-symmetry and marginal symmetry for multi-way tables," *Journal de la Société Française de Statistique*, vol. 148, pp. 3-36, 2007.
- [8] K. Tahata, H. Yamamoto, and S. Tomizawa, "Orthogonality of decompositions of symmetry into extended symmetry and marginal equipoment for multi-way tables with ordered categories," *Austrian Journal of Statistics*, vol. 37, pp. 185-194, 2008.
- [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] J. B. Lang, and A. Agresti, "Simultaneously modeling joint and marginal distributions of multivariate categorical responses," *Journal of the American Statistical Association*, vol. 89, pp. 625-632, 1994.
- [11] J. B. Lang, "On the partitioning of goodness-of-fit statistics for multivariate categorical response models," *Journal of the American Statistical Association*, vol. 91, pp. 1017-1023, 1996.
- [12] J. Aitchison, "Large-sample restricted parametric test," *Journal of the Royal Statistical Society, Series B*, vol. 24, pp. 234-250, 1962.
- [13] J. N. Darroch, and S. D. Silvey, "On testing more than one hypothesis," *Annals of Mathematical Statistics*, vol. 34, pp. 555-567, 1963.
- [14] K. Tahata, and S. Tomizawa, "Orthogonal decomposition of point-symmetry for multi-way tables," *Advances in Statistical Analysis*, vol. 92, pp. 255-269, 2008.
- [15] M. Haber, "Maximum likelihood methods for linear and log-linear models in categorical data," *Computational Statistics and Data Analysis*, vol. 3, pp. 1-10, 1985.
- [16] C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed. New York: Wiley, 1973.
- [17] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. AC-19, pp. 716-723, 1974.
- [18] S. Konishi, and G. Kitagawa, *Information Criteria and Statistical Modeling*, New York: Springer, 2008.
- [19] S. M. Miller, "Comparative social mobility," *Current Sociology*, vol. 9, pp. 1-61, 1960.

Received: August 20, 2010

Revised: September 15, 2010

Accepted: September 20, 2010

© Kurakami 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.