Journal of Mathematics and Statistics

Original Research Paper

Measure of Departure from Partial Symmetry for Square Contingency Tables

Yusuke Saigusa, Kouji Tahata and Sadao Tomizawa

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

Article history
Received: 23-05-2016
Revised: 02-08-2016
Accepted: 04-08-2016

Corresponding Author:
Yusuke Saigusa
Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan
Email: saigusaysk@gmail.com

Abstract: For square contingency tables, the present paper newly considers the partial symmetry model which indicates that there is a symmetric structure of probabilities for at least one of pairs of symmetric cells. It also proposes the measure to express the degree of departure from the partial symmetry model. Examples are given.

Keywords: Measure, Partial Symmetry, Square Contingency Table, Symmetry

Introduction

Consider a square $r\times r$ contingency table with the same row and column classifications. Let p_{ij} denote the cell probability that an observation falls in its (i, j) cell ($i = 1,\ldots,r$; $j = 1,\ldots,r$). Consider the Symmetry (S) model as follows:

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

where, $\psi_{ij} = \psi_{ji}$ for $i\neq j$ (Bowker, 1948; Bishop *et al*., 1975, p. 282).

For the analysis of data, the S model may fit the data poorly because it has the strong restriction. When the S model fits the data poorly, many statisticians may be interested in applying some models which have weaker restriction than the S model. There are some symmetry or asymmetry models; for instance, the marginal homogeneity model (Stuart, 1955), the quasi-symmetry model (Caussinus, 1965), the conditional symmetry model (McCullagh, 1978), the diagonals-parameter symmetry model (Goodman, 1979) and the cumulative diagonals-parameter symmetry model (Tomizawa, 1993; Tahata and Tomizawa, 2014), etc.

On the other hands, some statisticians may be interested in measuring the degree of departure from the S model when the model fits the data poorly.

Assume that $p_{ij} + p_{ji}>0$ for $i\neq j$. Let $p_{ij}^{*} = p_{ij} / \delta$ and $p_{ij}^{c} = p_{ij} / \left(p_{ij} + p_{ji}\right)$ for $i\neq j$ with $\delta = \Sigma\Sigma_{i\neq j}\,p_{ij}$. Tomizawa *et al.* (1998) gave the measure to express the degree of departure from the S model as follows:

$$\Phi_S^{(\lambda)} = \sum_{i=1}^{r-1}\sum_{j=i+1}^{r}(p_{ij}^{*} + p_{ji}^{*})\left[1-\frac{\lambda 2^{\lambda}}{2^{\lambda}-1}H_{ij}^{(\lambda)}\right] \quad \text{for } \lambda > -1,$$

where:

$$H_{ij}^{(\lambda)} = \frac{1}{\lambda}\left[1-\left(p_{ij}^{c}\right)^{\lambda+1}-\left(p_{ji}^{c}\right)^{\lambda+1}\right],$$

and the value at $\lambda = 0$ is taken to be the limit as $\lambda\to 0$, where λ is a real-valued parameter which is chosen by the user. Note that $H_{ij}^{(\lambda)}$ is the Patil and Taillie (1982) diversity index of degree λ which includes the Shannon entropy. We point out that $\Phi_S^{(\lambda)}$ is expressed as the weighted arithmetic mean of the diversity index $H_{ij}^{(\lambda)}$.

Table 1a is taken from Hashimoto (1999, p. 151) and Tables 1b and 1c are taken from Bishop *et al.* (1975, p. 100). These data relate father's and son's occupational status categories in Japan, in Denmark and in British. The smaller category number means the higher status in each of Tables 1a-c. For example, on the father's and son's occupational status in Denmark, it seems that there is a symmetric structure of probabilities for several pairs although there is not a symmetric structure of probabilities for all pairs.

Table 1. Cross-classifications of father's and son's occupational status (a) in Japan (Hashimoto, 1999, p.151), (b) in Denmark and (c) in British (Bishop *et al*., 1975, p.100)

(a) in Japan

Father's status	Son's status (1)	(2)	(3)	(4)	(5)	Total
(1)	39	39	39	57	23	197
(2)	12	78	23	23	37	173
(3)	6	16	78	23	20	143
(4)	18	80	79	126	31	334
(5)	28	106	136	122	628	1020
Total	103	319	355	351	739	1867

(b) in Denmark

Father's status	Son's status (1)	(2)	(3)	(4)	(5)	Total
(1)	18	17	16	4	2	57
(2)	24	105	109	59	21	318
(3)	23	84	289	217	95	708
(4)	8	49	175	348	198	778
(5)	6	8	69	201	246	530
Total	79	263	658	829	562	2391

(c) in British

Father's status	Son's status (1)	(2)	(3)	(4)	(5)	Total
(1)	50	45	8	18	8	129
(2)	28	174	84	154	55	495
(3)	11	78	110	223	96	518
(4)	14	150	185	714	447	1510
(5)	3	42	72	320	411	848
Total	106	489	459	1429	1017	3500

Now we consider the model expressed as:

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

where, $\psi_{st} = \psi_{ts}$ for at least one (s, t) with $s \neq t$. We shall refer to this model as the Partial Symmetry (PS) model. Since the S model indicates that p_{ij} equals p_{ji} for all (i, j), the PS model is implied by the S model.

For each of Tables 1a-c, the PS model means that the probability that a father's occupational status is i and son's occupational status is j, equals the probability that the father's occupational status is j and son's occupational status is i for at least one (i, j), $i = 1,\ldots,5$; $j = 1,\ldots,5$; $i \neq j$.

We are now interested in measuring the degree of departure from the PS model than the S model.

By the way, Tomizawa *et al*. (2004) gave the measure in the form of geometric mean, which describes the strength of association between the row and column variables for two-way contingency table, although the detail is omitted. In order to express the degree of departure from the PS model, we shall consider the geometric mean type measure.

In the present paper, section 2 proposes a new measure which expresses the degree of departure from the PS model. Section 3 gives the approximate confidence interval of the proposed measure. Section 4 gives Examples. Section 5 compares two measures and shows that the proposed measure is appropriate for measuring the degree of departure from PS. Section 6 presents concluding remarks.

Measure

Assume that $p_{ij} + p_{ji} > 0$ for $i \neq j$. Consider the measure defined by:

$$\Phi_P^{(\lambda)} = \prod_{i=1}^{r-1} \prod_{j=i+1}^{r} \left[1 - \frac{\lambda 2^{\lambda}}{2^{\lambda} - 1} H_{ij}^{(\lambda)} \right]^{\left(p_{ij}^{*} + p_{ji}^{*}\right)} \quad \text{for} \quad \lambda > -1,$$

where:

$$H_{ij}^{(\lambda)} = \frac{1}{\lambda}\left[1 - \left(p_{ij}^{c}\right)^{\lambda+1} - \left(p_{ji}^{c}\right)^{\lambda+1}\right],$$

and the value at $\lambda = 0$ is taken to be the limit as $\lambda \to 0$ and λ is a real-valued parameter which is chosen by the user. Note that $\Phi_P^{(\lambda)}$ is expressed as the weighted geometric mean of the diversity index. The measure $\Phi_P^{(\lambda)}$ must lie between 0 and 1 since $0 \le H_{ij}^{(\lambda)} \le \left(2^{\lambda} - 1\right) / \left(\lambda 2^{\lambda}\right)$ for $i<j$. For any $\lambda(>-1)$, (i) $\Phi_P^{(\lambda)}$ takes the minimum value 0 if and only if there is a structure of PS in the table and (ii) $\Phi_P^{(\lambda)}$ takes the maximum value 1 if and only if the degree of departure from PS is the largest in the sense that $p_{ij}^{c} = 1$ (then $p_{ji}^{c} = 0$) or $p_{ji}^{c} = 1$ (then $p_{ij}^{c} = 0$) for all (i, j), $i \neq j$.

It is easily seen that the value of $\Phi_P^{(\lambda)}$ is less than or equal to the value of $\Phi_S^{(\lambda)}$. It may be natural because the necessary and sufficient condition for $\Phi_P^{(\lambda)}$ taking the minimum value 0 is weaker than that for $\Phi_S^{(\lambda)}$ taking the minimum value 0, and the necessary and sufficient condition for $\Phi_P^{(\lambda)}$ taking the maximum value 1 is same as that for $\Phi_S^{(\lambda)}$ taking the maximum value 1.

We point out that $\Phi_P^{(\lambda)}$ is appropriate for the contingency table with the nominal categories, because the value of $\Phi_P^{(\lambda)}$ is invariant for same arbitrary permutations of the categories of rows and columns, namely, $\Phi_P^{(\lambda)}$ does not depend on the order of the categories.

Approximate Confidence Interval of Measure

Assume that a multinomial distribution applies to the $r \times r$ table. We shall obtain the approximate standard error and the large-sample confidence interval of $\Phi_P^{(\lambda)}$. Let n_{ij} denote the observed frequency of (i, j) cell in the table $(i = 1, \dots, r; j = 1, \dots, r)$, and let n denote the total number of observations, i.e., $n = \Sigma\Sigma n_{ij}$. The sample version of $\Phi_P^{(\lambda)}$, denoted by $\hat{\Phi}_P^{(\lambda)}$, is $\Phi_P^{(\lambda)}$ with (p_{ij}) replaced by $(\hat{p}_{ij})$, where $\hat{p}_{ij} = n_{ij} / n$. Using the delta method (Agresti, 2013, p.587), $\sqrt{n}\left(\hat{\Phi}_P^{(\lambda)} - \Phi_P^{(\lambda)}\right)$ has asymptotically (as $n \to \infty$) a normal distribution with mean zero and variance σ^2, where:

$$\sigma^2 = \sum_{i=1}^{r}\sum_{\substack{j=1\\ j\neq i}}^{r} p_{ij}\left(\Omega_{ij}^{(\lambda)}\right)^2 - \left(\sum_{i=1}^{r}\sum_{\substack{j=1\\ j\neq i}}^{r} p_{ij}\Omega_{ij}^{(\lambda)}\right)^2 \quad \text{for } \lambda > -1,$$

with:

$$\Omega_{ij}^{(\lambda)} = \frac{\Phi_P^{(\lambda)}}{\delta}\left(\log\omega_{ij}^{(\lambda)} + \frac{\xi_{ij}^{(\lambda)}}{\omega_{ij}^{(\lambda)}} - \sum_{k=1}^{r-1}\sum_{l=k+1}^{r}(p_{kl}^{*} + p_{lk}^{*})\log\omega_{kl}^{(\lambda)}\right),$$

$$\omega_{ij}^{(\lambda)} = \begin{cases} 1 - \dfrac{2^{\lambda}}{2^{\lambda}-1}\left(1 - \left(p_{ij}^{c}\right)^{\lambda+1} - \left(p_{ji}^{c}\right)^{\lambda+1}\right) & \text{for } \lambda \neq 0, \\ 1 - \dfrac{1}{\log 2}\left(-p_{ij}^{c}\log p_{ij}^{c} - p_{ji}^{c}\log p_{ji}^{c}\right) & \text{for } \lambda = 0, \end{cases}$$

$$\xi_{ij}^{(\lambda)} = \begin{cases} \dfrac{2^{\lambda}}{2^{\lambda}-1}(\lambda+1)\left(\left(p_{ij}^{c}\right)^{\lambda} p_{ji}^{c} - \left(p_{ji}^{c}\right)^{\lambda+1}\right) & \text{for } \lambda \neq 0, \\ \dfrac{1}{\log 2}\left(p_{ji}^{c}\log p_{ij}^{c} - p_{ji}^{c}\log p_{ji}^{c}\right) & \text{for } \lambda = 0. \end{cases}$$

We note that the asymptotic normal distribution of $\sqrt{n}\left(\hat{\Phi}_P^{(\lambda)} - \Phi_P^{(\lambda)}\right)$ is applicable only when $0 < \Phi_P^{(\lambda)} < 1$. Let $\hat{\sigma}^2$ be σ^2 with (p_{ij}) replaced by $(\hat{p}_{ij})$ The estimated approximate standard error of $\hat{\Phi}_P^{(\lambda)}$ is $\hat{\sigma}/\sqrt{n}$, and the approximate 100(1-α)% confidence interval of $\Phi_P^{(\lambda)}$ is $\hat{\Phi}_P^{(\lambda)} \pm z_{\alpha/2}\hat{\sigma}/\sqrt{n}$ where $z_{\alpha/2}$ is the quantile of the standard normal distribution corresponding to a two-tail probability equal to α.

Examples

Consider the data in Table 1 again. Tables 2 and 3 give the estimated values of measures $\Phi_S^{(\lambda)}$ and $\Phi_P^{(\lambda)}$ applied to each of Tables 1a-c. They also give the estimated approximate standard errors and the approximate 95% confidence intervals of the measures. From Table 3, for any λ(>-1), the confidence interval of $\Phi_P^{(\lambda)}$ applied to the data in Table 1a does not include 0. So there would not be the structure of PS in Table 1a. On the other hand, for any λ(>-1), the confidence intervals of $\Phi_P^{(\lambda)}$ applied to the data in each of Table 1b and 1c include 0. So there may be the structure of PS in each of Tables 1b and 1c.

We shall further compare the degrees of departure from PS for Tables 1a-c using $\Phi_P^{(\lambda)}$. Comparing the confidence intervals of $\Phi_P^{(\lambda)}$ for Tables 1a-c, for any λ(>-1), it is inferred that the degree of departure from PS for Table 1a is larger than that for each of Tables 1b and 1c. In a similar way, from Table 2, it is inferred that the degree of departure from S for Table 1a is larger than that for each of Tables 1b and 1c.

Table 2. The estimates of $\Phi_S^{(\lambda)}$, estimated approximate standard errors of $\hat{\Phi}_S^{(\lambda)}$ and approximate 95% confidence intervals of $\Phi_S^{(\lambda)}$, applied to each of Tables 1a-c

(a) Table 1a

λ	Estimated measure $\hat{\Phi}_S^{(\lambda)}$	Standard error	Confidence interval
-0.5	0.160	0.017	(0.127, 0.194)
0	0.252	0.025	(0.203, 0.301)
0.5	0.301	0.028	(0.245, 0.356)
1	0.323	0.030	(0.265, 0.381)
1.5	0.328	0.030	(0.270, 0.387)
2	0.323	0.030	(0.265, 0.381)
2.5	0.311	0.029	(0.254, 0.368)
3	0.295	0.029	(0.239, 0.351)

(b) Table 1b

λ	Estimated measure $\hat{\Phi}_S^{(\lambda)}$	Standard error	Confidence interval
-0.5	0.008	0.003	(0.002, 0.014)
0	0.013	0.005	(0.003, 0.023)
0.5	0.016	0.006	(0.004, 0.028)
1	0.018	0.007	(0.004, 0.031)
1.5	0.018	0.007	(0.004, 0.032)
2	0.018	0.007	(0.004, 0.031)
2.5	0.017	0.007	(0.004, 0.029)
3	0.015	0.006	(0.003, 0.027)

(c) Table 1c

λ	Estimated measure $\hat{\Phi}_S^{(\lambda)}$	Standard error	Confidence interval
-0.5	0.008	0.003	(0.003, 0.013)
0	0.013	0.004	(0.005, 0.022)
0.5	0.017	0.005	(0.006, 0.027)
1	0.018	0.006	(0.007, 0.030)
1.5	0.019	0.006	(0.007, 0.030)
2	0.018	0.006	(0.007, 0.030)
2.5	0.017	0.006	(0.006, 0.028)
3	0.016	0.005	(0.006, 0.026)

Table 3. The estimates of $\Phi_P^{(\lambda)}$, estimated approximate standard errors of $\hat{\Phi}_P^{(\lambda)}$ and approximate 95% confidence intervals of $\Phi_P^{(\lambda)}$, applied to each of Tables 1a, 1b and 1c

(a) Table 1a

λ	Estimated measure $\hat{\Phi}_P^{(\lambda)}$	Standard error	Confidence interval
-0.5	0.121	0.025	(0.072, 0.170)
0	0.193	0.039	(0.117, 0.270)
0.5	0.233	0.047	(0.142, 0.325)
1	0.252	0.050	(0.154, 0.350)
1.5	0.256	0.051	(0.157, 0.356)
2	0.252	0.050	(0.154, 0.350)
2.5	0.241	0.048	(0.147, 0.336)
3	0.227	0.046	(0.138, 0.317)

(b) Table 1b

λ	Estimated measure $\hat{\Phi}_P^{(\lambda)}$	Standard error	Confidence interval
-0.5	0.001	0.005	(-0.009, 0.012)
0	0.002	0.009	(-0.016, 0.020)
0.5	0.003	0.011	(-0.019, 0.025)
1	0.003	0.013	(-0.021, 0.028)
1.5	0.003	0.013	(-0.022, 0.029)
2	0.003	0.013	(-0.021, 0.028)
2.5	0.003	0.012	(-0.020, 0.026)
3	0.003	0.011	(-0.018, 0.024)

(c) Table 1c

λ	Estimated measure $\hat{\Phi}_P^{(\lambda)}$	Standard error	Confidence interval
-0.5	0.003	0.005	(-0.006, 0.013)
0	0.006	0.008	(-0.010, 0.021)
0.5	0.007	0.010	(-0.012, 0.027)
1	0.008	0.011	(-0.013, 0.029)
1.5	0.008	0.011	(-0.014, 0.030)
2	0.008	0.011	(-0.013, 0.029)
2.5	0.007	0.010	(-0.013, 0.028)
3	0.007	0.009	(-0.012, 0.025)

We point out that, for any λ(>-1) the estimated value of $\Phi_P^{(\lambda)}$ applied to each of Tables 1a-c is less than that of $\Phi_S^{(\lambda)}$.

Comparison between Measures

Consider the 4×4 artificial cell probability tables given in Table 4. There is a structure of PS in each of Tables 4a and 4b. Table 4c has a cell with probability zero in the lower left triangle. Also there are two or more cells with probabilities zeros in Tables 4d-h. Tables 5 and 6 shows the values of $\Phi_S^{(\lambda)}$ and $\Phi_P^{(\lambda)}$ applied to each table. We can see from Tables 4 and 6 that, for fixed λ, the value of $\Phi_P^{(\lambda)}$ increases as the number of cells with probabilities zeros in the 4×4 table increases. It may be natural to consider that the degree of departure from PS increases as the number of cells with probabilities zeros in the table increases.

Table 4. Artificial cell probability tables

(a)

0.140	0.017	0.033	0.018
0.017	0.141	0.004	0.018
0.066	0.016	0.140	0.015
0.054	0.090	0.090	0.141

(b)

0.159	0.016	0.038	0.012
0.016	0.159	0.008	0.011
0.076	0.032	0.160	0.009
0.036	0.055	0.054	0.159

(c)

0.164	0.012	0.039	0.015
0.000	0.165	0.013	0.010
0.078	0.052	0.164	0.004
0.045	0.050	0.024	0.165

(d)

0.158	0.071	0.008	0.013
0.000	0.158	0.020	0.018
0.000	0.080	0.159	0.004
0.039	0.090	0.024	0.158

(e)

0.194	0.042	0.004	0.021
0.000	0.195	0.009	0.016
0.000	0.036	0.194	0.002
0.000	0.080	0.012	0.195

(f)

0.202	0.033	0.023	0.005
0.000	0.202	0.008	0.018
0.000	0.000	0.202	0.002
0.000	0.090	0.012	0.203

(g)

0.216	0.001	0.008	0.025
0.000	0.217	0.020	0.010
0.000	0.000	0.217	0.010
0.000	0.000	0.060	0.216

(h)

0.215	0.079	0.013	0.021
0.000	0.215	0.004	0.018
0.000	0.000	0.215	0.005
0.000	0.000	0.000	0.215

Namely $\Phi_P^{(\lambda)}$ would be appropriate for measuring the degree of departure from PS. On the other hand, from Tables 4 and 5, $\Phi_S^{(\lambda)}$ is not appropriate for measuring the degree of departure from PS because, for fixed λ, the values of $\Phi_S^{(\lambda)}$ applied to Tables 4a and 4b are unequal (although there is a structure of PS in each of Tables 4a and 4b).

Table 5. Values of $\Phi_S^{(\lambda)}$ for Table 4

	λ		
Applied tables	0	0.5	1.5
Table 4a	0.246	0.293	0.320
Table 4b	0.216	0.258	0.282
Table 4c	0.244	0.286	0.310
Table 4d	0.452	0.498	0.523
Table 4e	0.535	0.579	0.604
Table 4f	0.589	0.631	0.654
Table 4g	0.691	0.728	0.748
Table 4h	1.000	1.000	1.000

Table 6. Values of $\Phi_P^{(\lambda)}$ for Table 4

	λ		
Applied tables	0	0.5	1.5
Table 4a	0.000	0.000	0.000
Table 4b	0.000	0.000	0.000
Table 4c	0.192	0.231	0.255
Table 4d	0.382	0.440	0.472
Table 4e	0.463	0.523	0.555
Table 4f	0.517	0.577	0.608
Table 4g	0.626	0.681	0.709
Table 4h	1.000	1.000	1.000

Concluding Remarks

For an $r \times r$ square contingency table, we have considered the PS model which has weaker restriction than the S model. The PS model indicates symmetry of probabilities for at least a pair of symmetric cells instead of all pairs of symmetric cells. We have proposed the measure to express the degree of departure from PS. The measure enables us to see how far cell probabilities are distant from those with a PS structure.

The readers may be interested in the relationship between the proposed measure and the goodness-of-fit test for the PS model. However it may be difficult to discuss the relationship.

We also have shown with Examples that $\Phi_P^{(\lambda)}$ is useful for expressing and comparing the degree of departure from the partial symmetry toward the complete asymmetry between different tables.

Acknowledgement

The authors thank the editor and referees for their helpful comments.

Author's Contributions

Yusuke Saigusa: Deriving the main results, Programming for examples and examinations and drafting and revising the paper.

Kouji Tahata: Deriving the main results, Programming for examples and examinations and checking the paper.

Sadao Tomizawa: Proposing the main idea and checking the paper.

Ethics

The authors declare that there is no conflict of interests regarding the publication of the present paper

References

Agresti, A., 2013. Categorical Data Analysis. 3rd Edn., John Wiley and Sons, Hoboken, N.J., ISBN-10: 1118710940, pp: 744.

Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1975. Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, ISBN-10: 0262520400, pp: 557.

Bowker, A.H., 1948. A test for symmetry in contingency tables. J. Amer. Statist. Assoc., 43: 572-574. DOI: 10.1080/01621459.1948.10483284

Caussinus, H., 1965. Contribution à l'analyse statistique des tableaux de corrélation. Ann. Fac. Sci. Univ. Toulouse, 29: 77-182. DOI: 10.5802/afst.519

Goodman, L.A., 1979. Multiplicative models for square contingency tables with ordered categories. Biometrika, 66: 413-418. DOI: 10.1093/biomet/66.3.413

Hashimoto, K., 1999. Gendai Nihon no Kaikyuu Kouzou (Class Structure in Modern Japan: Theory, Method and Quantitative Association). Toshindo Press, Tokyo.

McCullagh, P., 1978. A class of parametric models for the analysis of square contingency tables with ordered categories. Biometrika, 65: 413-418. DOI: 10.1093/biomet/65.2.413

Patil, G.P. and C. Taillie, 1982. Diversity as a concept and its measurement. J. Amer. Statist. Assoc., **77**: 548-561. DOI: 10.1080/01621459.1982.10477845

Stuart, A., 1955. A test for homogeneity of the marginal distributions in a two-way classification. Biometrika, 42: 412-416. DOI: 10.1093/biomet/42.3-4.412

Tahata, K. and S. Tomizawa, 2014. Symmetry and asymmetry models and decompositions of models for contingency tables. SUT J. Math., 50: 131-165.

Tomizawa, S., 1993. Diagonals-parameter symmetry model for cumulative probabilities in square contingency tables with ordered categories. Biometrics, 49: 883-887. DOI: 10.2307/2532209

Tomizawa, S., N. Miyamoto and H. Houya, 2004. Generalization of Cramer's coefficient of association for contingency tables. South Afr. Statist. J., 38: 1-24.

Tomizawa, S., T. Seo and H. Yamamoto, 1998. Power-divergence-type measure of departure from symmetry for square contingency tables that have nominal categories. J. Applied Stat., 25: 387-398. DOI: 10.1080/02664769823115