Measure of Symmetry with Minimum Variance for Square Contingency Tables

Kouji Yamamoto[1], Kouji Tahata[2], and Sadao Tomizawa[3]
[1] *Osaka University, Osaka, JAPAN*
[2,3] *Tokyo University of Science, Chiba, JAPAN*
[1] Corresponding author: Kouji Yamamoto,
e-mail: yamamoto-k@hp-crc.med.osaka-u.ac.jp

Abstract

For analyzing square contingency tables with same row and column classifications, models on symmetry instead of independence have been used. The simple symmetry model which represents the structure of symmetry with respect to the cell probabilities was introduced by Bowker. When the symmetry model does not hold for a given data, we are interested in measuring the degree of departure from symmetry. For square tables with nominal categories, a measure expressed by using Shannon entropy or Kullback-Leibler information was proposed by Tomizawa. The present paper proposes a measure for symmetry which is different from Tomizawa's measure, in the class of weighted averages. The proposed measure is an approximation to the measure in the class of weighted averages that has the smallest variance.

Keywords: categorical data, Kullback-Leibler information, Shannon entropy

1. Introduction

Consider an $r \times r$ square contingency table having same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the (i, j)th cell of the table $(i = 1, \ldots, r; j = 1, \ldots, r)$. For the analysis of such a square table, models on symmetry instead of independence have been used.

The simple symmetry model introduced by Bowker (1948) is defined by

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

This model indicates the structure of symmetry with respect to the cell probabilities.

When the symmetry model does not hold for a given data, we are interested in measuring the degree of departure from symmetry. Tomizawa (1994) proposed a measure to represent the degree of departure from symmetry expressed by using Shannon entropy or Kullback-Leibler information.

In the present paper, we propose a measure which represents the degree of departure from symmetry by a different approach from Tomizawa's measure, according to Agresti (1984, p. 170). Namely a new measure is expressed as the weighted average of sample submeasures for each of symmetric cells.

The purpose of the present paper is to propose a measure which is an approximation to the measure in the class of weighted averages that has the smallest variance.

2. The proposed measure

First, we shall introduce submeasures to represent the degree of departure from symmetry for (i, j)th cell and (j, i)th cell $(i < j)$.

Let n_{ij} denote the observed frequency in the (i, j)th cell of the table $(i = 1, \dots, r; j = 1, \dots, r)$ and $r_{ij} = n_{ij} + n_{ji}$. Also let π_{ij} denote the conditional probability that an observation will fall in the (i, j)th cell, $i \neq j$, of the table on condition that it falls in the (i, j)th cell or (j, i)th cell; namely $\pi_{ij} = p_{ij}/(p_{ij}+p_{ji})$. Then the probability mass function for $\{n_{ij}\}$, $i < j$, is the product of $r(r-1)/2$ binomials,

$$\prod_{i<j}\prod \frac{r_{ij}!}{n_{ij}!n_{ji}!}\pi_{ij}^{n_{ij}}\pi_{ji}^{n_{ji}}.$$

Then the symmetry model is also expressed as

$$\pi_{ij} = \pi_{ji} \left(= \frac{1}{2}\right) \quad (i < j).$$

We shall consider a submeasure to represent the degree of departure from symmetry for (i, j)th cell and (j, i)th cell $(i < j)$, as follows:

$$\hat{\phi}_{ij} = 1 - \frac{1}{\log 2}\hat{H}_{ij},$$

with

$$\hat{H}_{ij} = -\hat{\pi}_{ij}\log\hat{\pi}_{ij} - \hat{\pi}_{ji}\log\hat{\pi}_{ji},$$

where $\hat{\pi}_{ij} = n_{ij}/r_{ij}$. This is also expressed as

$$\hat{\phi}_{ij} = \frac{1}{\log 2}\hat{I}_{ij},$$

with

$$\hat{I}_{ij} = \hat{\pi}_{ij}\log\left(\frac{\hat{\pi}_{ij}}{1/2}\right) + \hat{\pi}_{ji}\log\left(\frac{\hat{\pi}_{ji}}{1/2}\right).$$

We note that ϕ_{ij} and H_{ij} (or I_{ij}) given by $\hat{\phi}_{ij}$ and $\hat{H}_{ij}$ (or $\hat{I}_{ij}$) with $\{\hat{\pi}_{ij}\}$ replaced by $\{\pi_{ij}\}$, are the population versions. Also note that the H_{ij} is Shannon entropy and I_{ij} is Kullback-Leibler information.

As a measure to represent the degree of departure from symmetry, we consider the class of weighted average of $\{\hat{\phi}_{ij}\}$; namely

$$\Phi = \sum_{i<j}\sum w_{ij}\hat{\phi}_{ij}, \tag{1}$$

where the weights $\{w_{ij}\}$ satisfying all $w_{ij} > 0$ and $\sum\sum w_{ij} = 1$.

Then, we consider the following measure which represents the degree of departure from symmetry:

$$\Phi_S = \sum_{i<j}\sum w_{ij}^*\hat{\phi}_{ij}, \tag{2}$$

where

$$w^*_{ij} = \frac{1/\sigma^2_{ij}}{\sum\sum_{s<t} 1/\sigma^2_{st}},$$

$$\sigma^2_{ij} = \frac{\pi_{ij}\pi_{ji}}{r_{ij}} \left(\frac{1}{\log 2} \left(\log \pi_{ij} - \log \pi_{ji} \right) \right)^2 \quad (i < j).$$

Note that the σ^2_{ij} $(i < j)$ is the variance of $\hat{\phi}_{ij}$ (see Appendix 1).

The measure (2) has the smallest variance among measures in the class of weighted averages. For the details, see Appendix 2.

So we shall propose the estimated measure as follows:

$$\hat{\Phi}_S = \sum\sum_{i<j} \hat{w}^*_{ij} \hat{\phi}_{ij},$$

where $\hat{w}^*_{ij}$ is given by w^*_{ij} with $\{\pi_{ij}\}$ replaced by $\{\hat{\pi}_{ij}\}$. This is an approximation of the measure in the class of weighted averages of $\{\hat{\phi}_{ij}\}$ that has the smallest variance.

3. Approximate standard error for the measure

We shall consider an approximate standard error for the measure Φ_S. We see from Appendix 1 that $\hat{\phi}_{ij}$ $(i < j)$ is asymptotically distributed normal as $N(\phi_{ij}, \sigma^2_{ij})$ independently. Thus the measure Φ_S is asymptotically distributed as normal with mean

$$\sum\sum_{i<j} w^*_{ij} \phi_{ij},$$

and variance

$$\sigma^2 = \sum\sum_{i<j} \left(w^*_{ij} \right)^2 \sigma^2_{ij}.$$

Let $\hat{\sigma}^2$ denote σ^2 with $\{\pi_{ij}\}$ replaced by $\{\hat{\pi}_{ij}\}$, i.e.,

$$\hat{\sigma}^2 = \sum\sum_{i<j} \left(\hat{w}^*_{ij} \right)^2 \hat{\sigma}^2_{ij}.$$

Therefore the estimated standard error for Φ_S is $\hat{\sigma}$.

Appendix 1

Let $\hat{\pi}'$ be the $1 \times r(r-1)$ vector

$$\hat{\pi}' = (\hat{\pi}'_{12}, \hat{\pi}'_{13}, \ldots, \hat{\pi}'_{r-1,r}),$$

where "$\prime$" denotes the transpose and

$$\hat{\pi}'_{ij} = (\hat{\pi}_{ij}, \hat{\pi}_{ji}) \quad (i < j).$$

Also, we define the vector π in terms of π_{ij}'s in a similar manner to $\hat{\pi}$. Then $\hat{\pi}$ is asymptotically distributed as normal $N(\pi, V(\pi))$, where $V(\pi)$ is the $r(r-1) \times r(r-1)$ matrix,

$$V(\pi) = \begin{pmatrix} V_{12}(\pi) & 0 & \cdots & 0 \\ 0 & V_{13}(\pi) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & V_{r-1,r}(\pi) \end{pmatrix},$$

where

$$V_{ij}(\pi) = \frac{1}{r_{ij}} \begin{pmatrix} \pi_{ij}(1-\pi_{ij}) & -\pi_{ij}\pi_{ji} \\ -\pi_{ij}\pi_{ji} & \pi_{ji}(1-\pi_{ji}) \end{pmatrix} \quad (i < j).$$

Then, let $\hat{\phi}$ be the $r(r-1)/2$ vector

$$\hat{\phi} = (\hat{\phi}_{12}, \hat{\phi}_{13}, \ldots, \hat{\phi}_{r-1,r})',$$

and we define the vector ϕ in terms of π_{ij}'s in a similar manner to $\hat{\phi}$. Using the delta method, $\hat{\phi}$ is asymptotically distributed as normal $N(\phi, \Sigma)$, where Σ is the $r(r-1)/2 \times r(r-1)/2$ matrix,

$$\begin{aligned} \Sigma &= \left(\frac{\partial\phi}{\partial\pi'}\right) V(\pi) \left(\frac{\partial\phi}{\partial\pi'}\right)' \\ &= \begin{pmatrix} \sigma_{12}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{13}^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{r-1,r}^2 \end{pmatrix}, \end{aligned}$$

where

$$\sigma_{ij}^2 = \frac{\pi_{ij}\pi_{ji}}{r_{ij}} \left(\frac{1}{\log 2}\left(\log \pi_{ij} - \log \pi_{ji}\right)\right)^2 \quad (i < j).$$

Appendix 2

Let w be the $r(r-1)/2 \times 1$ vector

$$w = (w_{12}, w_{13}, \ldots, w_{r-1,r})'.$$

Then the measure (1) is expressed as

$$\Phi = w'\hat{\phi},$$

where $\hat{\phi}$ is the $r(r-1)/2 \times 1$ vector defined in Appendix 1.

From Appendix 1, the expectation and variance of Φ is approximately calculated as follows:

$$E(\Phi) = w'\phi,$$
$$Var(\Phi) = w'\Sigma w.$$

Then, we can obtain the following w^* so as to minimize $Var(\Phi)$ with the constrain that $w'1_d$ ($d = r(r-1)/2$) is unity (1_d is the $d \times 1$ vector of 1 elements):

$$w^* = \frac{1}{1_d' \Sigma^{-1} 1_d} \Sigma^{-1} 1_d.$$

We see from Appendix 1 that the w^*_{ij} ($i < j$) is expressed as follows:

$$w^*_{ij} = \frac{1/\sigma^2_{ij}}{\sum\sum_{s<t} 1/\sigma^2_{st}}.$$

References

Agresti, A. (1984). *Analysis of Ordinal Categorical Data.* Wiley, New York.

Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association*, **43**, 572-574.

Tomizawa, S. (1994). Two kinds of measures of departure from symmetry in square contingency tables having nominal categories. *Statistica Sinica*, **4**, 325-334.