Measures on Proportional Reduction in Error by Arithmetic, Geometric and Harmonic Means for Multi-way Contingency Tables

Kouji Yamamoto (Corresponding author)

Center for Clinical Investigation and Research, Osaka University Hospital

2-15, Yamadaoka, Suita 565-0871, Osaka, Japan

E-mail: yamamoto-k@hp-crc.med.osaka-u.ac.jp

Yuri Nozaki

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science

2641, Yamazaki, Noda City, Chiba 278-8510, Japan

E-mail: yr.nozaki@gmail.com

Sadao Tomizawa

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science

2641, Yamazaki, Noda City, Chiba 278-8510, Japan

E-mail: tomizawa@is.noda.tus.ac.jp

Received: December 27, 2011 Accepted: January 11, 2012 Published: May 1, 2012

doi:10.5539/ijsp.v1n1p36 URL: http://dx.doi.org/10.5539/ijsp.v1n1p36

Abstract

For multi-way contingency tables with nominal categories, this paper proposes three kinds of proportional reduction in error measures, which describe the relative decrease in the probability of making an error in predicting the value of one variable when the values of the other variables are known, as opposed to when they are not known. The measures have forms of arithmetic, geometric and harmonic means. An example is shown.

Keywords: Arithmetic mean, Geometric mean, Harmonic mean, Proportional reduction in error

1. Introduction

Consider an $R \times C$ contingency table with both nominal categories of the explanatory variable X and the response variable Y. Let p_{ij} denote the probability that an observation will fall in the ith category of X and in the jth category of Y ($i = 1, \ldots, R$; $j = 1, \ldots, C$). Goodman and Kruskal (1954) proposed two kinds of measures, i.e., (1) the measure which describes the proportional reduction in variation (PRV) in predicting the Y category obtained when the X category is known, as opposed to when the X category is not known, and (2) the measure which describes the proportional reduction in error (PRE) in predicting it. Although the details are omitted, some PRV measures are considered by, e.g., Theil (1970), Tomizawa, Seo and Ebi (1997), Tomizawa and Ebi (1998), Tomizawa and Yukawa (2003), and Yamamoto, Miyamoto and Tomizawa (2010).

The present paper considers the PRE measures. Goodman and Kruskal (1954) proposed the PRE measure as

$$\lambda_B = \frac{(1 - p_{\bullet m_0}) - \sum_{i=1}^{R} p_{i\bullet}\left(1 - \left(\frac{p_{im_i}}{p_{i\bullet}}\right)\right)}{1 - p_{\bullet m_0}} = \frac{\sum_{i=1}^{R} p_{im_i} - p_{\bullet m_0}}{1 - p_{\bullet m_0}},$$

where

$$p_{im_i} = \max_j(p_{ij}), \quad p_{\bullet m_0} = \max_j(p_{\bullet j}), \quad p_{i\bullet} = \sum_{t=1}^{C} p_{it}, \quad p_{\bullet j} = \sum_{s=1}^{R} p_{sj};$$

also see Bishop, Fienberg and Holland (1975, p. 388), and Everitt (1992, p. 58). This measure describes the relative decrease in the probability of making an error in predicting the value of Y when the value of X is known, as opposed to when it is not known. The measure λ_B has the properties that (i) $0 \leq \lambda_B \leq 1$, (ii) $\lambda_B = 0$ if and only if the information about the explanatory variable X does not reduce the probability of making an error in predicting the category of the variable Y, and (iii) $\lambda_B = 1$ if and only if no error is made, given knowledge of the explanatory variable X; namely there is complete predictive association.

Next, consider the reverse case which is the explanatory variable Y and the response variable X. The following measure λ_A is suitable for predictions of X from Y, defined by

$$\lambda_A = \frac{\sum_{j=1}^{C} p_{M_j j} - p_{M_0 \bullet}}{1 - p_{M_0 \bullet}},$$

where

$$p_{M_j j} = \max_i (p_{ij}), \quad p_{M_0 \bullet} = \max_i (p_{i \bullet});$$

see Goodman and Kruskal (1954).

The measures λ_B and λ_A are specifically designed for the situation in which the explanatory and response variables are defined. Now consider the situation where the explanatory and response variables are not defined. In this case, the following measure λ is given:

$$\lambda = \frac{\sum_{i=1}^{R} p_{i m_i} + \sum_{j=1}^{C} p_{M_j j} - p_{\bullet m_0} - p_{M_0 \bullet}}{2 - p_{\bullet m_0} - p_{M_0 \bullet}};$$

see Goodman and Kruskal (1954). This indicates the PRE in predicting the category of either variable as between knowing and not knowing the category of the other variable. Also, the measure λ is the weighted sum of the measures λ_B and λ_A.

For a two-way contingency table with both nominal categories, Yamamoto and Tomizawa (2010) proposed new PRE measures, say Λ, expressed as the arithmetic, geometric and harmonic means of λ_B and λ_A. For a two-way contingency table with nominal-ordinal categories, Yamamoto, Nozaki and Tomizawa (2011) proposed a PRE measure although the detail is omitted.

The purpose of the present paper is to extend the Yamamoto and Tomizawa's (2010) measures into T-way contingency tables ($T \geq 3$) with all nominal categories. Section 2 proposes measures for three-way tables ($T = 3$), and Section 3 extends them for multi-way ($T \geq 4$) and expresses as more generalized form including such three kinds of means. Section 4 analyzes data as an example.

2. New PRE Measures for Three-way Contingency Tables

2.1 Measures

Consider an $R \times C \times L$ contingency table with variables X, Y and Z which have all nominal categories. Let p_{ijk} denote the probability of that an observation will fall in the (i, j, k)th cell of the table ($i = 1, \ldots, R; j = 1, \ldots, C; k = 1, \ldots, L$). When the explanatory and response variables are not defined, namely, we cannot specifically define which of the variables is a response, we consider three kinds of prediction, predicting X, predicting Y and predicting Z.

First, consider the table with a response variable X and two explanatory variables Y and Z. In this case, a PRE measure, which describes the relative decrease in the probability of making error in predicting the value of X when the values of the other variables, Y and Z, are known, as opposed to when they are not known, is defined by

$$\lambda_A^{(3)} = \frac{\sum_{j=1}^{C} \sum_{k=1}^{L} p_{m_{jk} jk} - p_{m_1 \bullet\bullet}}{1 - p_{m_1 \bullet\bullet}},$$

where

$$p_{m_{jk} jk} = \max_i (p_{ijk}), \quad p_{m_1 \bullet\bullet} = \max_i (p_{i \bullet\bullet}), \quad p_{i \bullet\bullet} = \sum_{t=1}^{C} \sum_{u=1}^{L} p_{itu}.$$

Similarly, each PRE measure for the table as having a response variable Y and two explanatory variables X and Z and as having a response variable Z and two explanatory variables X and Y is defined by

$$\lambda_B^{(3)} = \frac{\sum_{i=1}^{R}\sum_{k=1}^{L} p_{im_{ik}k} - p_{\bullet m_2 \bullet}}{1 - p_{\bullet m_2 \bullet}},$$

and

$$\lambda_C^{(3)} = \frac{\sum_{i=1}^{R}\sum_{j=1}^{C} p_{ijm_{ij}} - p_{\bullet\bullet m_3}}{1 - p_{\bullet\bullet m_3}},$$

where

$$p_{im_{ik}k} = \max_j (p_{ijk}), \quad p_{\bullet m_2 \bullet} = \max_j (p_{\bullet j \bullet}), \quad p_{\bullet j \bullet} = \sum_{s=1}^{R}\sum_{u=1}^{L} p_{sju},$$

$$p_{ijm_{ij}} = \max_k (p_{ijk}), \quad p_{\bullet\bullet m_3} = \max_k (p_{\bullet\bullet k}), \quad p_{\bullet\bullet k} = \sum_{s=1}^{R}\sum_{t=1}^{C} p_{stk}.$$

Then, we shall propose three kinds of new PRE measures as follows:

$$\lambda_a^{(3)} = \frac{\lambda_A^{(3)} + \lambda_B^{(3)} + \lambda_C^{(3)}}{3},$$

$$\lambda_g^{(3)} = \sqrt[3]{\lambda_A^{(3)}\lambda_B^{(3)}\lambda_C^{(3)}},$$

and

$$\lambda_h^{(3)} = \frac{3}{\dfrac{1}{\lambda_A^{(3)}} + \dfrac{1}{\lambda_B^{(3)}} + \dfrac{1}{\lambda_C^{(3)}}}.$$

The measures $\lambda_a^{(3)}$, $\lambda_g^{(3)}$ and $\lambda_h^{(3)}$ are the arithmetic mean, geometric mean and harmonic mean of the $\lambda_A^{(3)}$, $\lambda_B^{(3)}$ and $\lambda_C^{(3)}$, respectively.

Let λ^* denote each of measures $\lambda_a^{(3)}$, $\lambda_g^{(3)}$ and $\lambda_h^{(3)}$. Each measure has the properties that (i) λ^* must lie between 0 and 1, (ii) $\lambda^* = 0$ if and only if the information about two variables does not reduce the probability of making an error in predicting the category of the other variable, and (iii) $\lambda^* = 1$ if and only if no error is made, given knowledge of two variables; namely there is complete predictive association. We point out that if the variables are independent, then the measure λ^* takes 0, but the converse need not hold. Note that when the values of $\lambda_A^{(3)}$, $\lambda_B^{(3)}$ and $\lambda_C^{(3)}$ are 0 such as the variables are independent, the measure $\lambda_h^{(3)}$ cannot measure the PRE. So in such a case, the measures $\lambda_a^{(3)}$ and $\lambda_g^{(3)}$ should be used as a PRE measure.

We see that

$$\min\left(\lambda_A^{(3)}, \lambda_B^{(3)}, \lambda_C^{(3)}\right) \le \lambda_h^{(3)} \le \lambda_g^{(3)} \le \lambda_a^{(3)} \le \max\left(\lambda_A^{(3)}, \lambda_B^{(3)}, \lambda_C^{(3)}\right),$$

where the equality holds if and only if $\lambda_A^{(3)} = \lambda_B^{(3)} = \lambda_C^{(3)}$.

2.2 Approximate Confidence Interval for Measures

Let n_{ijk} denote the observed frequency in the (i, j, k)th cell of the table ($i = 1, \ldots, R; j = 1, \ldots, C; k = 1, \ldots, L$). Assuming that $\{n_{ijk}\}$ result from full multinomial sampling, we consider an approximate standard error and large-sample confidence interval for λ^*, using the delta method, descriptions of which are given by Bishop et al. (1975, Sec. 14.6). The sample version of λ^*, i.e., $\hat{\lambda}^*$, is given by λ^* with $\{p_{ijk}\}$ replaced by $\{\hat{p}_{ijk}\}$, where $\hat{p}_{ijk} = n_{ijk}/n$ and $n = \sum\sum n_{ijk}$. Using the delta method, $\sqrt{n}(\hat{\lambda}^* - \lambda^*)$ has asymptotically (as $n \to \infty$) a normal distribution with mean 0 and variance $\sigma^2[\lambda^*]$, where

$$\sigma^2[\lambda^*] = \sum_{i=1}^{R}\sum_{j=1}^{C}\sum_{k=1}^{L}\left(\frac{\partial \lambda^*}{\partial p_{ijk}}\right)^2 p_{ijk} - \left(\sum_{s=1}^{R}\sum_{t=1}^{C}\sum_{u=1}^{L}\left(\frac{\partial \lambda^*}{\partial p_{stu}}\right) p_{stu}\right)^2.$$

For measures $\lambda_a^{(3)}$, $\lambda_g^{(3)}$ and $\lambda_h^{(3)}$, the variances are

$$\text{(a)}\ \sigma^2\left[\lambda_a^{(3)}\right] = \frac{1}{9}\left[\sum_{i=1}^{R}\sum_{j=1}^{C}\sum_{k=1}^{L}(U_{ijk})^2 p_{ijk} - \left(\sum_{s=1}^{R}\sum_{t=1}^{C}\sum_{u=1}^{L} U_{stu}p_{stu}\right)^2\right],$$

$$\text{(b)}\ \sigma^2\left[\lambda_g^{(3)}\right] = \frac{1}{9}\left(\lambda_g^{(3)}\right)^{-4}\left[\sum_{i=1}^{R}\sum_{j=1}^{C}\sum_{k=1}^{L}(V_{ijk})^2 p_{ijk} - \left(\sum_{s=1}^{R}\sum_{t=1}^{C}\sum_{u=1}^{L} V_{stu}p_{stu}\right)^2\right],$$

$$\text{(c)}\ \sigma^2\left[\lambda_h^{(3)}\right] = \frac{1}{9}\left(\lambda_h^{(3)}\right)^{4}\left[\sum_{i=1}^{R}\sum_{j=1}^{C}\sum_{k=1}^{L}(W_{ijk})^2 p_{ijk} - \left(\sum_{s=1}^{R}\sum_{t=1}^{C}\sum_{u=1}^{L} W_{stu}p_{stu}\right)^2\right],$$

where

$$\begin{aligned} U_{ijk} &= \Delta_{ijk(1)} + \Delta_{ijk(2)} + \Delta_{ijk(3)}, \\ V_{ijk} &= \Delta_{ijk(1)}\lambda_B^{(3)}\lambda_C^{(3)} + \Delta_{ijk(2)}\lambda_A^{(3)}\lambda_C^{(3)} + \Delta_{ijk(3)}\lambda_A^{(3)}\lambda_B^{(3)}, \\ W_{ijk} &= \frac{\Delta_{ijk(1)}}{\left(\lambda_A^{(3)}\right)^2} + \frac{\Delta_{ijk(2)}}{\left(\lambda_B^{(3)}\right)^2} + \frac{\Delta_{ijk(3)}}{\left(\lambda_C^{(3)}\right)^2}, \end{aligned}$$

with

$$\Delta_{ijk(1)} = \frac{I(i = m_{jk})(1 - p_{m_1\bullet\bullet}) - I(i = m_1)\left(1 - \sum_{j=1}^{C}\sum_{k=1}^{L} p_{m_{jk}\bullet\bullet}\right)}{(1 - p_{m_1\bullet\bullet})^2},$$

$$\Delta_{ijk(2)} = \frac{I(j = m_{ik})(1 - p_{\bullet m_2\bullet}) - I(j = m_2)\left(1 - \sum_{i=1}^{R}\sum_{k=1}^{L} p_{\bullet m_{ik}\bullet}\right)}{(1 - p_{\bullet m_2\bullet})^2},$$

$$\Delta_{ijk(3)} = \frac{I(k = m_{ij})(1 - p_{\bullet\bullet m_3}) - I(k = m_3)\left(1 - \sum_{i=1}^{R}\sum_{j=1}^{C} p_{\bullet\bullet m_{ij}}\right)}{(1 - p_{\bullet\bullet m_3})^2},$$

and $I(\cdot)$ is the indicator function.

Let $\hat{\sigma}^2[\lambda^*]$ denote $\sigma^2[\lambda^*]$ with $\{p_{ijk}\}$ replaced by $\{\hat{p}_{ijk}\}$. Then, $\hat{\sigma}[\lambda^*]/\sqrt{n}$ is an estimated standard error for $\hat{\lambda}^*$, and $\hat{\lambda}^* \pm z_{1-\alpha/2}\hat{\sigma}[\lambda^*]/\sqrt{n}$ is an approximate $100(1-\alpha)\%$ confidence interval for λ^*, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution.

3. Extension to Multi-way Contingency Tables

3.1 Measures

Consider an $R_1 \times R_2 \times \cdots \times R_T$ contingency table with nominal categories in which the $(T-1)$ explanatory variables and one response variable are not defined. Let $p_{i_1 i_2 \cdots i_T}$ denote the probability that an observation will fall in the $(i_1, i_2, \cdots, i_T)$th cell of the table ($i_k = 1, \ldots, R_k$; $k = 1, \ldots, T$) and X_k ($k = 1, \cdots, T$) denote the kth variable. For $k = 1, \ldots, T$, a PRE measure in predicting the value of X_k is defined by

$$\lambda_k^{(T)} = \frac{\sum_{i_1=1}^{R_1}\cdots\sum_{i_{k-1}=1}^{R_{k-1}}\sum_{i_{k+1}=1}^{R_{k+1}}\cdots\sum_{i_T=1}^{R_T} p^{(k)}_{m_{i_1\cdots i_{k-1}i_{k+1}\cdots i_T}} - p^{(k)}_{m_k}}{1 - p^{(k)}_{m_k}},$$

where

$$p^{(k)}_{m_{i_1\cdots i_{k-1}i_{k+1}\cdots i_T}} = \max_{i_k}(p_{i_1\cdots i_k\cdots i_T}), \quad p^{(k)}_{m_k} = \max_{i_k}\left(p^{(k)}_{i_k}\right),$$

and $p^{(k)}_{i_k} = \mathrm{P}(X_k = i_k)$. Then, we shall extend the measures as follows:

$$\lambda_a^{(T)} = \frac{1}{T}\sum_{k=1}^{T}\lambda_k^{(T)},$$

$$\lambda_g^{(T)} = \sqrt[T]{\prod_{k=1}^{T} \lambda_k^{(T)}},$$

and

$$\lambda_h^{(T)} = \frac{T}{\sum_{k=1}^{T} \frac{1}{\lambda_k^{(T)}}}.$$

The measures $\lambda_a^{(T)}$, $\lambda_g^{(T)}$ and $\lambda_h^{(T)}$ are the arithmetic mean, geometric mean and harmonic mean of the $\lambda_1^{(T)}$ through $\lambda_T^{(T)}$, respectively.

Let $\Lambda^{(T)}$ denote each of measures $\lambda_a^{(T)}$, $\lambda_g^{(T)}$ and $\lambda_h^{(T)}$. Each measure has the properties that (i) $\Lambda^{(T)}$ must lie between 0 and 1, (ii) $\Lambda^{(T)} = 0$ if and only if the information about $(T-1)$ variables does not reduce the probability of making an error in predicting the category of the other variable, and (iii) $\Lambda^{(T)} = 1$ if and only if no error is made, given knowledge of $(T-1)$ variables; namely there is complete predictive association. We point out that if all variables are independent, then the measure $\Lambda^{(T)}$ takes 0, but the converse need not hold. Note that when $\lambda_k^{(T)} = 0$ $(k = 1, \cdots T)$ such as all variables are independent, the measure $\lambda_h^{(T)}$ cannot measure the PRE. So in such a case, the measures $\lambda_a^{(T)}$ and $\lambda_g^{(T)}$ should be used as a PRE measure.

We see that

$$\min\left(\lambda_1^{(T)}, \cdots, \lambda_T^{(T)}\right) \le \lambda_h^{(T)} \le \lambda_g^{(T)} \le \lambda_a^{(T)} \le \max\left(\lambda_1^{(T)}, \cdots, \lambda_T^{(T)}\right),$$

where the equality holds if and only if $\lambda_1^{(T)}$ through $\lambda_T^{(T)}$ are all equal.

We note that $\Lambda^{(T)}$ when $T = 2$ is equivalent to the measure Λ proposed in Yamamoto et al. (2010).

3.2 Generalization of the Measures

Considering the monotonic function g, we shall propose a generalized measure, which includes the measures $\lambda_a^{(T)}$, $\lambda_g^{(T)}$ and $\lambda_h^{(T)}$, as follows:

$$\Lambda^{(T)} = g^{-1}\left(\frac{\sum_{k=1}^{T} g\left(\lambda_k^{(T)}\right)}{T}\right).$$

The functions g and g^{-1} are differentiable functions. Especially, (i) when $g(x) = x$, the measure $\Lambda^{(T)}$ is identical to $\lambda_a^{(T)}$, (ii) when $g(x) = \log x$, the measure $\Lambda^{(T)}$ is identical to $\lambda_g^{(T)}$, and (iii) when $g(x) = 1/x$, the measure $\Lambda^{(T)}$ is identical to $\lambda_h^{(T)}$.

$\Lambda^{(T)}$ has the same properties as $\lambda_a^{(T)}$, $\lambda_g^{(T)}$ and $\lambda_h^{(T)}$ (see Section 3.1).

3.3 Approximate Confidence Interval for Measures

Let $n_{i_1 i_2 \cdots i_T}$ denote the observed frequency in the $(i_1, i_2, \ldots, i_T)$th cell of the table $(i_k = 1, \ldots, R_k; k = 1, \ldots, T)$. Assume that a multinomial distribution applies to the $R_1 \times R_2 \times \cdots \times R_T$ table. In a similar way to the case of $T = 3$, $\sqrt{n}(\hat{\Lambda}^{(T)} - \Lambda^{(T)})$ (n is sample size and $\hat{\Lambda}^{(T)}$ is the estimated measure) has asymptotically a normal distribution with mean 0 and variance

$$\sigma^2\left[\Lambda^{(T)}\right] = \sum_{j_1=1}^{R_1} \cdots \sum_{j_T=1}^{R_T} \left(\frac{\partial \Lambda^{(T)}}{\partial p_{j_1 \cdots j_T}}\right)^2 p_{j_1 \cdots j_T} - \left(\sum_{s_1=1}^{R_1} \cdots \sum_{s_T=1}^{R_T} \left(\frac{\partial \Lambda^{(T)}}{\partial p_{s_1 \cdots s_T}}\right) p_{s_1 \cdots s_T}\right)^2.$$

For measures $\lambda_a^{(T)}$, $\lambda_g^{(T)}$ and $\lambda_h^{(T)}$, the variances are

$$\text{(a)}\ \sigma^2\left[\lambda_a^{(T)}\right] = \frac{1}{T^2}\left[\sum_{j_1=1}^{R_1} \cdots \sum_{j_T=1}^{R_T} (U_{j_1 \cdots j_T})^2 p_{j_1 \cdots j_T} - \left(\sum_{s_1=1}^{R_1} \cdots \sum_{s_T=1}^{R_T} U_{s_1 \cdots s_T} p_{s_1 \cdots s_T}\right)^2\right],$$

$$\text{(b)}\ \sigma^2\left[\lambda_g^{(T)}\right] = \frac{1}{T^2}\left(\lambda_g^{(T)}\right)^{2(1-T)}\left[\sum_{j_1=1}^{R_1} \cdots \sum_{j_T=1}^{R_T} (V_{j_1 \cdots j_T})^2 p_{j_1 \cdots j_T} - \left(\sum_{s_1=1}^{R_1} \cdots \sum_{s_T=1}^{R_T} V_{s_1 \cdots s_T} p_{s_1 \cdots s_T}\right)^2\right],$$

$$\text{(c)}\ \sigma^2\left[\lambda_h^{(T)}\right] = \frac{1}{T^2}\left(\lambda_h^{(T)}\right)^{4}\left[\sum_{j_1=1}^{R_1} \cdots \sum_{j_T=1}^{R_T} (W_{j_1 \cdots j_T})^2 p_{j_1 \cdots j_T} - \left(\sum_{s_1=1}^{R_1} \cdots \sum_{s_T=1}^{R_T} W_{s_1 \cdots s_T} p_{s_1 \cdots s_T}\right)^2\right],$$

where

$$U_{j_1\cdots j_T} = \sum_{k=1}^{T} \Delta_{j_1\cdots j_T(k)},$$
$$V_{j_1\cdots j_T} = \sum_{k=1}^{T} \Delta_{j_1\cdots j_T(k)} \lambda_1^{(T)} \cdots \lambda_{k-1}^{(T)} \lambda_{k+1}^{(T)} \cdots \lambda_T^{(T)},$$
$$W_{j_1\cdots j_T} = \sum_{k=1}^{T} \frac{\Delta_{j_1\cdots j_T(k)}}{\left(\lambda_k^{(T)}\right)^2},$$

with

$$\Delta_{j_1\cdots j_T(k)} = \frac{1}{\left(1 - p_{m_k}^{(k)}\right)^2} \left[I(j_k = m_{j_1\cdots j_{k-1} j_{k+1}\cdots j_T}) \left(1 - p_{m_k}^{(k)}\right) - I(j_k = m_k) \left(1 - \sum_{i_1=1}^{R_1} \cdots \sum_{i_{k-1}=1}^{R_{k-1}} \sum_{i_{k+1}=1}^{R_{k+1}} \cdots \sum_{i_T=1}^{R_T} p_{m_{i_1\cdots i_{k-1} i_{k+1}\cdots i_T}}^{(k)} \right) \right],$$

and $I(\cdot)$ is the indicator function.

Then, we can construct an asymptotic confidence interval using estimated variance although the detail is omitted.

4. An Example

Consider the data in Table 1, taken from Goodman (1975), which shows the McHugh test data on creative ability in machine design. This table cross-classifies 137 engineers with respect to their dichotomized scores (above the subtest mean (1) or below the subtest mean (2)) obtained on each of four different subtests that were supposed to measure creative ability in machine design. There are sixteen response patterns because the table has four variables (items A, B, C and D) and each has two categories.

Now, we are interested in what degree the relative decrease in the probability of making an error in predicting the value of one variable when we know the values of the other three variables as opposed to when we do not know them is. We shall analyze these data by using the proposed measure because the explanatory and response variables are not defined. When we use the measure $\lambda_a^{(4)}$, for example, the estimated value of $\lambda_a^{(4)}$ is 0.470 (Table 2). We see that in prediction of one of the variables from the others, the information reduces the probability of making an error by 47.0%. Similarly, the estimated values of $\lambda_g^{(4)}$ and $\lambda_h^{(4)}$ are 0.469 and 0.467, respectively. So we can also obtain a similar interpretation for the data.

We are also interested in the values of test statistic for the hypotheses of independence of (1) item A and items (B, C, D), (2) B and (A, C, D), (3) C and (A, B, D), and (4) D and (A, B, C). The values of Pearson's chi-squared statistic are 35.93 for (1), 37.67 for (2), 48.17 for (3), and 42.06 for (4) with seven degrees of freedom. Therefore, we can see the strong association between one of the variables and the other three variables. So, it would be meaningful to see the values of proposed measures.

5. Concluding Remarks

For analyzing multi-way (T-way) contingency tables with nominal categories, we have proposed three kinds of PRE measures which describes the relative decrease in the probability of making error in predicting value of one variable when the values of the other variables are known, as opposed to when they are not known. The proposed measures include arithmetic mean ($\lambda_a^{(T)}$), geometric mean ($\lambda_g^{(T)}$) and harmonic mean ($\lambda_h^{(T)}$). These measures are useful for analyzing the table which explanatory and response variables are not defined. A point to notice is that the measure $\lambda_h^{(T)}$ cannot measure the PRE when the variables are independent and/or any $\lambda_k^{(T)}$ ($k = 1, \ldots, T$) is 0. In such a case, the measures $\lambda_a^{(T)}$ and $\lambda_g^{(T)}$ should be used. It is difficult to discuss how to choose between three propositions: arithmetic, geometric or harmonic mean. We recommend the use of $\lambda_a^{(T)}$ for the simple interpretation.

In addition, the measure $\Lambda^{(T)}$, including $\lambda_a^{(T)}$, $\lambda_g^{(T)}$ and $\lambda_h^{(T)}$, is invariant under arbitrary permutations of the categories. Therefore the measure is suitable for analyzing the data on a nominal scale, but it is possible for analyzing the data on an ordinal scale because it only requires a categorical scale.

Acknowledgment

The authors would like to thank an editor and a referee for the helpful comments and suggestions.

References

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Massachusetts: The MIT Press.

Everitt, B. S. (1992). *The Analysis of Contingency Tables*. (2nd ed.). London: Chapman and Hall.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*, 732-764.

Goodman, L. A. (1975). A new model for scaling response patterns: An application of the quasi-independence concept. *Journal of the American Statistical Association, 70*, 755-768.

Theil, H. (1970). On the estimation of relationships involving qualitative variables. *American Journal of Sociology, 76*, 103-154. http://dx.doi.org/10.1086/224909

Tomizawa, S., & Ebi, M. (1998). Generalized proportional reduction in variation measure for multi-way contingency tables. *Journal of Statistical Research, 32*, 75-84.

Tomizawa, S., Seo, T., & Ebi, M. (1997). Generalized proportional reduction in variation measure for two-way contingency tables. *Behaviormetrika, 24*, 193-201. http://dx.doi.org/10.2333/bhmk.24.193

Tomizawa, S., & Yukawa, T. (2003). Proportional reduction in variation measures of departure from cumulative dichotomous independence for square contingency tables with same ordinal classifications. *Far East Journal of Theoretical Statistics, 11*, 133-165.

Yamamoto, K., Miyamoto, N., & Tomizawa, S. (2010). Harmonic, geometric and arithmetic means type uncertainty measures for two-way contingency tables with nominal categories. *Advances and Applications in Statistics, 17*, 143-159.

Yamamoto, K., Nozaki, Y., & Tomizawa, S. (2011). On measure of proportional reduction in error for nominal-ordinal contingency tables. *Journal of Statistics and Management Systems, 14*, 767-773.

Yamamoto, K., & Tomizawa, S. (2010). Measures of proportional reduction in error for two-way contingency tables with nominal categories. *Biostatistics, Bioinformatics and Biomathematics, 2*, 43-52.

Table 1. Frequency of occurrence of response patterns for the four machine design subtests (Goodman, 1975)

Response pattern				
item				Observed
A	B	C	D	frequencies
1	1	1	1	23
1	1	1	2	5
1	1	2	1	5
1	1	2	2	14
1	2	1	1	8
1	2	1	2	2
1	2	2	1	3
1	2	2	2	8
2	1	1	1	6
2	1	1	2	3
2	1	2	1	2
2	1	2	2	4
2	2	1	1	9
2	2	1	2	3
2	2	2	1	8
2	2	2	2	34
Total				137

Table 2. Estimates of the measures, approximate standard errors for them and approximate 95% confidence intervals for the measures, applied to Table 1

Measures	Estimated measure	Standard error	Confidence interval
$\lambda_a^{(4)}$	0.470	0.065	(0.343, 0.597)
$\lambda_g^{(4)}$	0.469	0.066	(0.340, 0.597)
$\lambda_h^{(4)}$	0.467	0.066	(0.337, 0.598)