İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 11, 2018, 2, 93-108

Geliş/Received:09.07.2018, Kabul/Accepted: 12.10.2018

www.istatistikciler.org

Derleme Makalesi / Compilation Article

Measures of departure from marginal homogeneity model in square contingency tables

Gökçen Altun
Department of Statistics
Bartin University
74100, Bartin, Turkey
gokcenefendioglu@gmail.com
0000-0003-4311-6508

Serpil Aktaş
Department of Statistics
Hacettepe University
06800, Beytepe, Ankara, Turkey
spxl@hacettepe.edu.tr
0000-0003-3364-6388

Abstract

In this study, the measures representing the departure from MH model is introduced for square contingency tables and standard errors of these measures is calculated. These measures lies between (-1,1) and represent the departure degree from MH model. Three real data sets are used as the illustrative examples. While the measures of departure are calculated for each category, the categories causing the departure from the MH model can be obtained and interpreted.

Keywords: Square contingency tables, marginal homogeneity model, departure measure

Öz

Karesel olumsallık tablolarında marjinal homojenlik modelinden sapma ölçüleri

Bu çalışmada, karesel olumsallık tabloları için MH modelinden ayrılışı ifade eden ölçümler ifade edilmiş ve bu ölçümlerin standart hataları hesaplanmıştır. Çalışmadaki sapma ölçüsü (-1,1) aralığında değer alır ve MH modelinden ayrılış derecesini ifade eder. Örnek olarak üç gerçek veri kümesi kullanılmıştır. Her bir kategori için sapma ölçüsü hesaplanmış, MH modelinden sapmaya neden olan kategoriler elde edilerek yorumlanmıştır.

Anahtar sözcükler: *Karesel olumsallık tablosu, marjinal homojenlik modeli, sapma ölçüsü*

1. Introduction

Square contingency tables that arise in dependent samples where the row and column variables have same level. Some specific models used in the analysis of these kinds of tables. The Marginal Homogeneity Model (MH) is one of them [1]. The model assumes that the marginal totals are symmetric. This model indicates that the row marginal distribution is identical with column marginal distribution. Consider an $R \times R$ square contingency table with ordered categories. Let p_{ij}

denotes the probability that an observation fall in the ith row and jth column of the table $(i, j = 1, 2, \ldots, R)$, and also let X and Y denote the row and column variables, respectively. The MH [2,3] model is defined by;

$$F_i^X = F_i^Y, \quad i = 1, 2, \ldots, R-1 \tag{1}$$

where,

$$F_i^X = \Pr(X \le i) = \sum_{k=1}^{i} p_{k.} \qquad i = 1, 2, \ldots, R-1$$
$$F_i^Y = \Pr(Y \le i) = \sum_{k=1}^{i} p_{.k} \qquad i = 1, 2, \ldots, R-1$$

(2)

with

$$p_{k.} = \sum_{t=1}^{R} p_{kt} \quad , \quad p_{.k} = \sum_{s=1}^{R} p_{sk} \quad , \quad i = 1, 2, \ldots, R$$

(3)

Let F_i^X and F_i^Y denote the cumulative marginal probability of X and Y, respectively, and also let $p_{k.}$ and $p_{.k}$ denote the marginal probability of X and Y, respectively. This model has $(R-1)$ degrees of freedom [4]. Let T_i^X and T_i^Y denote the conditional cumulative marginal probabilities of X and Y, respectively. Then the MH model may be expressed as; $T_i^X = T_i^Y, \quad i = 1, 2, \ldots, R-1$

where,

$$T_i^X = \Pr(X \le i \mid \mathrm{X} \ne \mathrm{Y}) = \sum_{k=1}^{i} p_{k.}^c \qquad i = 1, 2, \ldots, R-1$$
$$T_i^Y = \Pr(Y \le i \mid \mathrm{X} \ne \mathrm{Y}) = \sum_{k=1}^{i} p_{.k}^c \qquad i = 1, 2, \ldots, R-1$$

(4)

with

$$p_{k.}^{c}=\frac{1}{\delta}\left(p_{k.}-p_{kk}\right)\quad,\quad p_{.k}^{c}=\frac{1}{\delta}\left(p_{.k}-p_{kk}\right)\quad,\quad \delta=\sum\sum_{s\neq t}p_{st}$$

(5)

When the MH model does not hold, we are interested in considering the measures to represent the degree of departure from MH model [5]. The deviation measure, which refers to the deviation from the model as a percentage, is used in particular to compare the data sets with the same properties obtained at different time intervals.

2. Measures of departure from marginal homogeneity model

Tahata (2012) proposed a measure which represents the degree of departure from MH model. The measure is calculated over the observed frequencies gives the degree of the departure from MH. The proposed measure lies between -1 to 1. Zero represents the MH model holds and, 1 represents the degree of departure from marginal homogeneity is maximum. According to different values of the measures, upper-left-marginal inhomogeneity and lower-left-marginal inhomogeneity concept are reviewed in [6]. We will give these measures in the following subsections.

2.1. Measure I

Measure I is calculated using the marginal probabilities. This measure is the arithmetic mean of two sub-measures.

Sub-measure 2.1.1.

Let

$$\Delta_1=\sum_{i=1}^{R-1}\left(F_i^X+F_i^Y\right) \tag{6}$$

and

$$F_{1(i)}^{*}=\frac{F_i^X}{\Delta_1},\ \ F_{2(i)}^{*}=\frac{F_i^Y}{\Delta_1},\ \ \text{for}\ \ i=1,\dots,R-1 \tag{7}$$

Assuming that $\left\{F_i^X + F_i^Y \neq 0\right\}$, consider the sub-measure is defined as,

$$\Psi_1 = \frac{4}{\pi}\sum_{i=1}^{R-1}\left(F_{1(i)}^* + F_{2(i)}^*\right)\left(\theta_i^{(1)} - \frac{\pi}{4}\right) \tag{8}$$

where

$$\theta_i^{(1)} = \sin^{-1}\left(\frac{F_i^Y}{\sqrt{\left(F_i^X\right)^2 + \left(F_i^Y\right)^2}}\right) \tag{9}$$

Noting that, $\theta_i^{(1)}$ is between 0 and $\pi/2$, we can conclude that;

i. $-1 \le \Psi_1 \le 1$, ii. $\Psi_1 = -1$ if and only if $F_i^Y = 0$ and $F_i^X > 0$ $i = 1,\dots,R-1$, iii. $\Psi_1 = 1$ if and only if $F_i^X = 0$ and $F_i^Y > 0$ $i = 1,\dots,R-1$, iv. When the MH model holds, $\Psi_1 = 0$.

Sub-measure 2.1.2.
Let

$$S_i^X = 1 - F_i^X,\; S_i^Y = 1 - F_i^Y \text{ for } i = 1,\dots,R-1 \tag{10}$$

The MH model can be expressed as,

$$S_i^X = S_i^Y \text{ for } i = 1,\dots,R-1 \tag{11}$$

Let

$$\Delta_2 = \sum_{i=1}^{R-1}\left(S_i^X + S_i^Y\right) \tag{12}$$

$$S_{1(i)}^* = \frac{S_i^X}{\Delta_2},\; S_{2(i)}^* = \frac{S_i^Y}{\Delta_2}, \text{ for } i = 1,\dots,R-1 \tag{13}$$

Assuming that $\left\{S_i^X + S_i^Y \neq 0\right\}$, consider the sub-measure is defined as

$$\Psi_2 = \frac{4}{\pi}\sum_{i=1}^{R-1}\left(S_{1(i)}^{*} + S_{2(i)}^{*}\right)\left(\theta_i^{(2)} - \frac{\pi}{4}\right) \tag{14}$$

where

$$\theta_i^{(2)} = \sin^{-1}\left(\frac{S_i^X}{\sqrt{\left(S_i^X\right)^2 + \left(S_i^Y\right)^2}}\right) \tag{15}$$

It can be said that $\theta_i^{(2)}$ is between 0 and $\pi/2$ and,

i. $-1 \le \Psi_2 \le 1$, ii. $\Psi_2 = -1$ if and only if $S_i^X = 0$ and $S_i^Y > 0$ $i = 1,...,R-1$, iii. $\Psi_2 = 1$ if and only if $S_i^Y = 0$ and $S_i^X > 0$ $i = 1,...,R-1$. When the MH model holds, $\Psi_2 = 0$

Complete measure 1

The complete measure is calculated from Equation (16);

$$\Psi = \frac{\Psi_1 + \Psi_2}{2} \tag{16}$$

When $\Psi = 0$, we shall refer to this structure as the MH. As the measure Ψ approaches to -1, the departure from MH would be greater toward the upper-right-marginal inhomogeneity. While Ψ approaches 1, it becomes greater toward the lower-left-marginal inhomogeneity. Upper-right marginal inhomogeneity and lower-left marginal inhomogeneity for 4x4 square contingency table are represented in Figure 1 and Figure 2.

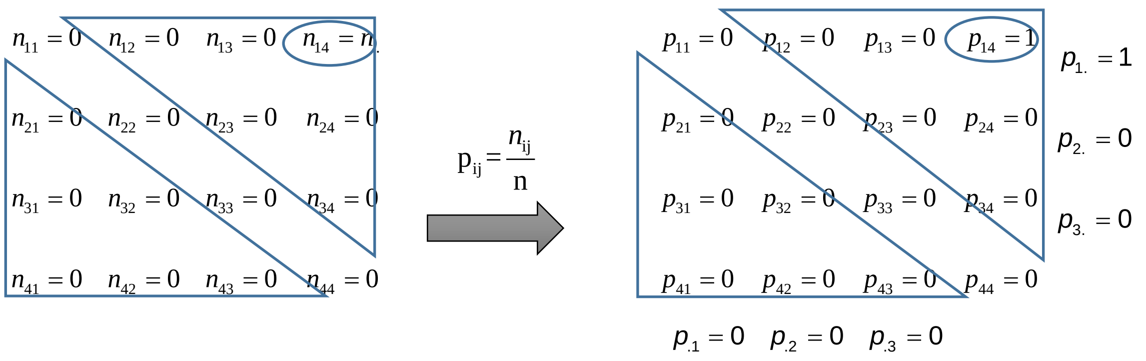

Figure 1. Upper-right marginal inhomogeneity for 4x4 square contingency table

The cumulative probabilities given in the Equation (1) can be calculated for the table in Figure 1 in terms of the marginal probabilities as follows:

$$\left.\begin{array}{l} \left.\begin{array}{l} F_1^X = p_{1.} = 1 \\ F_2^X = p_{1.} + p_{2.} = 1 \\ F_3^X = p_{1.} + p_{2.} + p_{3.} = 1 \end{array}\right\} \quad F_i^X = 1 \\ \\ \left.\begin{array}{l} F_1^Y = p_{.1} = 0 \\ F_2^Y = p_{.1} + p_{.2} = 0 \\ F_3^Y = p_{.1} + p_{.2} + p_{.3} = 0 \end{array}\right\} \quad F_i^Y = 0 \end{array}\right\} \quad \Psi = -1$$

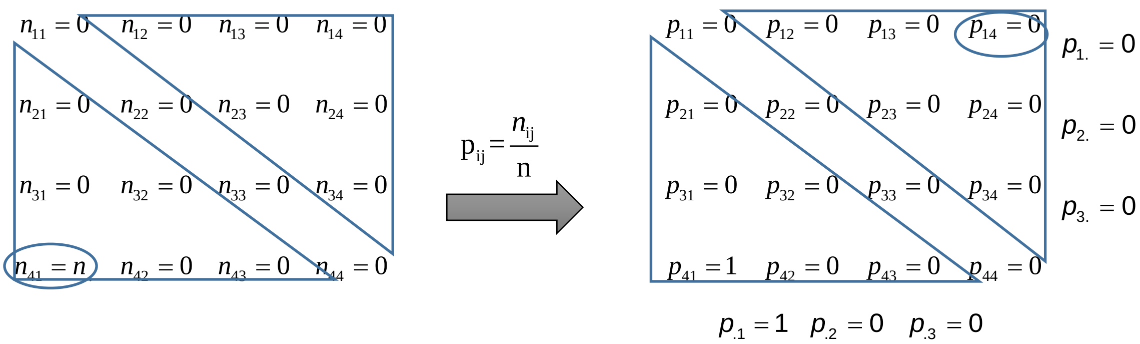

Figure 2. Lower-left marginal inhomogeneity for 4x4 square contingency table

The cumulative probabilities given in the Equation (1) can be calculated for the table in Figure 2 in terms of the marginal probabilities as follows:

$$\left.\begin{array}{l}\left.\begin{array}{l} F_1^X = p_{1.} = 0 \\ F_2^X = p_{1.} + p_{2.} = 0 \\ F_3^X = p_{1.} + p_{2.} + p_{3.} = 0 \end{array}\right\} \quad F_i^X = 0 \\ \left.\begin{array}{l} F_1^Y = p_{.1} = 1 \\ F_2^Y = p_{.1} + p_{.2} = 1 \\ F_3^Y = p_{.1} + p_{.2} + p_{.3} = 1 \end{array}\right\} \quad F_i^Y = 1 \end{array}\right\} \quad \Psi = 1$$

From Figure 1 and Figure 2 we can say that $\Psi = -1$ indicates that $p_{1R} = 1$ and the other cell probabilities are zero , $\Psi = 1$ indicates that $p_{R1} = 1$ and the other cell probabilities are zero.

2.2. Measure II

Measure II is calculated using the marginal conditional probabilities. This measure is mean of two sub-measures.

Sub-measure 2.2.1.

Let $\Delta_3 = \sum_{i=1}^{R-1} \left(T_i^X + T_i^Y \right)$

(17)

and

$$T_{1(i)}^* = \frac{T_i^X}{\Delta_3}, \quad T_{2(i)}^* = \frac{T_i^Y}{\Delta_3}, \quad \text{for } i = 1, \ldots, R-1$$

(18)

Assuming that $\left\{ T_i^X + T_i^Y \neq 0 \right\}$, consider the sub-measure is defined as,

$$\Upsilon_1 = \frac{4}{\pi}\sum_{i=1}^{R-1}\left(T^*_{1(i)} + T^*_{2(i)}\right)\left(\theta_i^{(3)} - \frac{\pi}{4}\right)$$

(19)

where,

$$\theta_i^{(3)} = \sin^{-1}\left(\frac{T_i^Y}{\sqrt{\left(T_i^X\right)^2 + \left(T_i^Y\right)^2}}\right) . \qquad (20)$$

Sub-measure 2.2.2.

Let $U_i^X = 1 - T_i^X, \;\; U_i^Y = 1 - U_i^Y \;$ for $\; i = 1,...,R-1$

(21)

The MH model alternatively can be expressed as,

$U_i^X = U_i^Y \;$ for $\; i = 1,...,R-1$

(22)

Let $\Delta_4 = \sum_{i=1}^{R-1}\left(U_i^X + U_i^Y\right)$

(23)

and

$U^*_{1(i)} = \frac{U_i^X}{\Delta_4}, \;\; U^*_{2(i)} = \frac{U_i^Y}{\Delta_4}, \;$ for $\; i = 1,...,R-1$

(24)

Assuming that $\left\{U_i^X + U_i^Y \neq 0\right\}$, consider the sub-measure is defined as,

$$\Upsilon_2 = \frac{4}{\pi}\sum_{i=1}^{R-1}\left(U^*_{1(i)} + U^*_{2(i)}\right)\left(\theta_i^{(4)} - \frac{\pi}{4}\right)$$

(25)

where

$$\theta_i^{(4)} = \sin^{-1}\left(\frac{U_i^X}{\sqrt{\left(U_i^X\right)^2 + \left(U_i^Y\right)^2}}\right)$$

(26)

Complete measure 2

Consider a complete measure defined by;

$$\Upsilon = \frac{\Upsilon_1 + \Upsilon_2}{2}$$

(27)

When $\Upsilon = 0$, we shall refer to this structure as the MH. Similarly as in the previous case, as the measure Υ approaches to -1, the departure from MH would be greater toward the upper-right conditional marginal inhomogeneity. While Υ approaches to 1, it becomes greater toward the conditional lower-left-marginal inhomogeneity.

3. Approximate confidence interval for the measures

Let n_{ij} denote the observed frequency in the ith row and jth column of the table $(i = 1, \dots, R; j = 1, \dots, R)$. Assuming that a multinomial distribution is applied to the $R \times R$ table, we shall consider an approximate standard error and large-sample confidence interval for the measure Ψ.

Delta method;

In statistics, the **delta method** is a result concerning the approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator. The delta method is defined as

$$\sqrt{n}[X_n - \theta] \to N(0, \sigma^2)$$

where X_n is a sequence of random variables, θ and σ^2 are finite valued constants. Using delta method, we obtain the following theorems.

Theorem 1: $\sqrt{n}[\widehat{\Psi} - \Psi]$ has asymptotically (as $n \to \infty$) a normal distribution with mean zero and variance, where $\sigma^2(\widehat{\Psi})$

$$\sigma^2\left[\hat{\Psi}\right]=\frac{1}{4}\sum_{i=1}^{R}\sum_{j=1}^{R}\left(a_{ij}+b_{ij}\right)^2 p_{ij}$$

(28)

with

$$a_{ij}=\frac{4}{\pi\Delta_1}\sum_{k=1}^{R-1}\left[\left\{I\left(i\le k\right)+I\left(j\le k\right)\right\}\theta_k^{(1)}+\frac{F_k^X+F_k^Y}{\left(F_k^X\right)^2+\left(F_k^Y\right)^2}\left\{-I\left(i\le k\right)F_k^Y+I\left(j\le k\right)F_k^X\right\}\right]-\frac{\left\{2R-\left(i+j\right)\right\}\left(\Psi_1+1\right)}{\Delta_1}$$

(29)

$$b_{ij}=\frac{4}{\pi\Delta_2}\sum_{k=1}^{R-1}\left[\left\{I\left(i>k\right)+I\left(j>k\right)\right\}\theta_k^{(2)}+\frac{S_k^X+S_k^Y}{\left(S_k^X\right)^2+\left(S_k^Y\right)^2}\left\{I\left(i>k\right)S_k^Y-I\left(j>k\right)S_k^X\right\}\right]-\frac{\left\{\left(i+j\right)-2\right\}\left(\Psi_2+1\right)}{\Delta_2}$$

$I(.)$ is the indicator function where $I(.)=1$ if the statement is true, 0 if it is not.

Theorem 2 : $\sqrt{n}\left[\widehat{\Upsilon}-\Upsilon\right]$ has asymptotically (as $n\to\infty$) a normal distribution with mean zero and variance, where $\sigma^2(\widehat{\Upsilon})$

$$\sigma^2\left[\hat{\Upsilon}\right]=\frac{1}{4}\sum_{i=1}^{R}\sum_{\substack{j=1\\ j\ne i}}^{R}\left(c_{ij}+d_{ij}\right)^2 p_{ij}$$

(30)

with

$$c_{ij}=\frac{4}{\delta\pi\Delta_3}\sum_{k=1}^{R-1}\left[\begin{array}{l}\left\{\left(I(i\le k)-T_k^X\right)+\left(I(j\le k)-T_k^Y\right)\right\}\left(\theta_k^{(3)}-\dfrac{\pi}{4}\right)\\ +\dfrac{T_k^X+T_k^Y}{\left(T_k^X\right)^2+\left(T_k^Y\right)^2}\left\{-\left(I(i\le k)-T_k^X\right)T_k^Y+\left(I(j\le k)-T_k^Y\right)T_k^X\right\}\end{array}\right]$$

$$-\frac{1}{\delta\Delta_3}\sum_{k=1}^{R-1}\left[\left(I(i\le k)-T_k^X\right)+\left(I(j\le k)-T_k^Y\right)\right]\Upsilon_1$$

$$d_{ij}=\frac{4}{\delta\pi\Delta_4}\sum_{k=1}^{R-1}\left[\begin{array}{l}\left\{\left(I(i>k)-U_k^X\right)+\left(I(j>k)-U_k^Y\right)\right\}\left(\theta_k^{(4)}-\dfrac{\pi}{4}\right)\\ +\dfrac{U_k^X+U_k^Y}{\left(U_k^X\right)^2+\left(U_k^Y\right)^2}\left\{\left(I(i>k)-U_k^X\right)T_k^Y-\left(I(j>k)-U_k^Y\right)U_k^X\right\}\end{array}\right]$$

$$-\frac{1}{\delta\Delta_4}\sum_{k=1}^{R-1}\left[\left(I(i>k)-U_k^X\right)+\left(I(j>k)-U_k^Y\right)\right]\Upsilon_2$$

(31)

Let $\hat{\sigma}^2(\widehat{\Psi})$ denote $\sigma^2(\widehat{\Psi})$ with p_{ij} replaced by $\hat{p}_{ij}$. Thus, the square root of $\hat{\sigma}^2(\widehat{\Psi})\ /n$ is an estimated standard error of $\widehat{\Psi}$ and the equation below

$$\hat{\Psi}\pm Z_{\alpha/2}\sqrt{\hat{\sigma}^2\left(\hat{\Psi}\right)/n}$$

is an approximate $100(1-\alpha)\%$ confidence interval for Ψ, where $Z_{\alpha/2}$ is the percentage point of the standard normal distribution corresponding to a two-tail probability of α. We also obtain the similar result for measure Υ [7].

4. Numerical example

As an illustrative example, distribution of spouses by respective educational level 1991, 2001, 2011 marriage and divorce data set in Turkey is used (Table 1-3). Data set, is organized according to four educational levels and sorted by from the low education level to higher as: (1) primary school, (2) middle school graduate, (3) high school graduate, (4) university graduate [8].

Table 1: Educational level of spouses in 1991

		Educational level of wife			
		1	2	3	4
Educational level of husband	1	11257	306	107	9
	2	2823	3284	217	12
	3	701	701	2723	65
	4	94	80	388	472

Table 2: Educational level of spouses in 2001

		Educational level of wife			
		1	2	3	4
Educational level of	1	15556	629	364	35
	2	3841	6379	585	25
	3	1360	5401	10141	250
	4	112	110	822	1529

Table 3: Educational level of spouses in 2011

		Educational level of wife			
		1	2	3	4
Educational level of husband	1	26644	2319	6220	881
	2	5672	1793	3164	532
	3	9518	2243	13736	4295
	4	1753	541	5733	9553

The MH model is applied to data sets and the results are given in Table 4.

Table 4: Likelihood ratio (G^2), degree of freedom, p-value for MH model

	df	G^2	p
1991	3	3272.61	<0.001
2001	3	7414.17	<0.001
2011	3	2822.35	<0.001

It is clear from Table 4 that the model does not fit the data sets. Since the MH model does not hold for the data sets, the next step would be the calculation the measures of departure from MH model. The measure of departure for each consecutive year are calculated and given below in detail:

Calculation for 1991;

$F_1^X = 0.503$, $F_2^X = 0.776$, $F_3^X = 0.956$, $F_1^Y = 0.640$, $F_2^Y = 0.828$, $F_3^Y = 0.976$, $\Delta_1 = 4.679$
$S_1^X = 0.497$, $S_2^X = 0.224$, $S_3^X = 0.044$, $S_1^Y = 0.360$, $S_2^Y = 0.172$, $S_3^Y = 0.027$, $\Delta_2 = 1.321$
$F_{1(1)}^* = 0.108$, $F_{1(2)}^* = 0.166$, $F_{1(3)}^* = 0.204$, $F_{2(1)}^* = 0.137$, $F_{2(2)}^* = 0.177$, $F_{2(3)}^* = 0.209$
$S_{1(1)}^* = 0.376$, $S_{1(2)}^* = 0.170$, $S_{1(3)}^* = 0.033$, $S_{2(1)}^* = 0.273$, $S_{2(2)}^* = 0.130$, $S_{2(3)}^* = 0.018$
$\theta_1^{(1)} = 57.57$, $\theta_2^{(1)} = 58.65$, $\theta_3^{(1)} = 46.55$, $\Psi_1 = 0.056$
$\theta_1^{(2)} = 54.48$, $\theta_2^{(2)} = 52.48$, $\theta_3^{(2)} = 61.39$, $\Psi_2 = 0.197$
$a_{11} = 0.015$, $a_{12} = -0.009$, $a_{13} = -0.011$, $a_{14} = -0.007$, $a_{21} = 0.0001$, $a_{22} = -0.025$, $a_{23} = -0.026$, $a_{24} = -0.0222$, $a_{31} = 0.008$, $a_{32} = -0.017$, $a_{33} = -0.018$, $a_{34} = -0.014$, $a_{41} = 0.022$, $a_{42} = -0.003$, $a_{43} = 0.004$, $a_{44} = 0$, $b_{11} = 0$, $b_{12} = 0.015$, $b_{13} = -0.059$, $b_{14} = 0.046$, $b_{21} = 0.016$, $b_{22} = 0.023$, $b_{23} = -0.043$, $b_{24} = 0.062$, $b_{31} = 0.006$, $b_{32} = -0.01$, $b_{33} = -0.054$, $b_{34} = 0.07$, $b_{41} = 0.143$, $b_{42} = 0.126$, $b_{43} = 0.082$, $b_{44} = 0.188$

Table 5: Estimate of Ψ , estimated approximate standard error for $\hat{\Psi}$, and approximate 95% confidence interval for Ψ, applied to Table 1

Table	$\hat{\Psi}$	S.E.	C.I.
1	0.1265	0.000144871	(0.1262 , 0.1268)

The results show that, the degree of departure from MH in 1991 is estimated as 12.65 percent of the maximum departure toward the lower-left-marginal inhomogeneity. All values in confidence interval for Ψ are positive. Therefore Ψ is statistically significant and, the structure of MH model for educational level of man and woman departs toward the lower-left-marginal inhomogeneity.

Table 6: Measures of without each groups for 1991

1991	Ψ
First group removed	0.0921
Second group removed	**0.0834**
Third group removed	0.1711
Fourth group removed	0.1592

When the second group (middle school) is removed from the table, departure is the smallest. This means that the departure from MH model is largest for the middle school graduates.

Calculations for 2001;

$F_1^X = 0.350$, $F_2^X = 0.580$, $F_3^X = 0.945$, $F_1^Y = 0.444$, $F_2^Y = 0.710$, $F_3^Y = 0.961$, $\Delta_1 = 3.99$
$S_1^X = 0.650$, $S_2^X = 0.420$, $S_3^X = 0.055$, $S_1^Y = 0.556$, $S_2^Y = 0.290$, $S_3^Y = 0.039$, $\Delta_2 = 2.01$
$F_{1(1)}^* = 0.088$, $F_{1(2)}^* = 0.145$, $F_{1(3)}^* = 0.237$, $F_{2(1)}^* = 0.111$, $F_{2(2)}^* = 0.178$, $F_{2(3)}^* = 0.241$
$S_{1(1)}^* = 0.323$, $S_{1(2)}^* = 0.209$, $S_{1(3)}^* = 0.027$, $S_{2(1)}^* = 0.277$, $S_{2(2)}^* = 0.144$, $S_{2(3)}^* = 0.019$
$\theta_1^{(1)} = 51.75$, $\theta_2^{(1)} = 50.76$, $\theta_3^{(1)} = 45.48$, $\Psi_1 = 0.076$
$\theta_1^{(2)} = 49.46$, $\theta_2^{(2)} = 55.38$, $\theta_3^{(2)} = 54.66$, $\Psi_2 = 0.15$
$a_{11} = 0.023$, $a_{12} = 0.004$, $a_{13} = -0.014$, $a_{14} = -0.0001$, $a_{21} = 0.015$, $a_{22} = -0.008$, $a_{23} = -0.026$, $a_{24} = -0.015$, $a_{31} = 0.008$, $a_{32} = -0.015$, $a_{33} = -0.033$, $a_{34} = -0.021$,

$a_{41} = 0.03$, $a_{42} = 0.007$, $a_{43} = -0.007$, $a_{44} = 0$, $b_{11} = 0$, $b_{12} = -0.037$, $b_{13} = -0.01$, $b_{14} = 0.01$, $b_{21} = -0.015$, $b_{22} = -0.052$, $b_{23} = -0.025$, $b_{24} = -0.025$, $b_{31} = 0.034$, $b_{32} = -0.004$, $b_{33} = 0.024$, $b_{34} = 0.043$, $b_{41} = 0.075$, $b_{42} = 0.037$, $b_{43} = 0.065$, $b_{44} = 0.085$.

Table 7: Estimate of Ψ , estimated approximate standard error for $\widehat{\Psi}$, and approximate 95% confidence interval for Ψ, applied to Table 2

Table	$\widehat{\Psi}$	S.E.	C.I.
2	0.113	0.000080642	(0.1128 , 0.1132)

We can see from this measure that the degree of departure from MH for the data in Table 2 is estimated to be 11.3 percent of the maximum departure toward the lower-left-marginal inhomogeneity. All values in confidence interval for Ψ are positive. Therefore, Ψ is statistically significant and, the structure of MH model for educational level of man and woman departs toward the lower-left-marginal inhomogeneity. Next step will be analyzing the levels. Each level is removed from the data and Ψ is recalculated (Table 8).

Table 8: Measures of without each groups for 2001

2001	Ψ
First group removed	0.1511
Second group removed	**0.0497**
Third group removed	0.1153
Fourth group removed	0.1480

When the second group (middle school) is removed from the table, departure is the smallest. This means that the departure from MH model is largest for the middle school graduates.

Calculations for 2011;

$F_1^X = 0.381$, $F_2^X = 0499$, $F_3^X = 0.814$, $F_1^Y = 0.461$, $F_2^Y = 0.534$, $F_3^Y = 0.839$, $\Delta_1 = 3.53$
$S_1^X = 0.619$, $S_2^X = 0.501$, $S_3^X = 0.186$, $S_1^Y = 0.539$, $S_2^Y = 0.466$, $S_3^Y = 0.161$, $\Delta_2 = 2.47$
$F_{1(1)}^* = 0.108$, $F_{1(2)}^* = 0.141$, $F_{1(3)}^* = 0.231$, $F_{2(1)}^* = 0.131$, $F_{2(2)}^* = 0.151$, $F_{2(3)}^* = 0.238$
$S_{1(1)}^* = 0.250$, $S_{1(2)}^* = 0.203$, $S_{1(3)}^* = 0.075$, $S_{2(1)}^* = 0.218$, $S_{2(2)}^* = 0.189$, $S_{2(3)}^* = 0.065$
$\theta_1^{(1)} = 50.43$, $\theta_2^{(1)} = 46.94$, $\theta_3^{(1)} = 45.87$, $\Psi_1 = 0.0499$
$\theta_1^{(2)} = 48.95$, $\theta_2^{(2)} = 47.07$, $\theta_3^{(2)} = 49.12$, $\Psi_2 = 0.071$
$a_{11} = 0.016$, $a_{12} = -0.009$, $a_{13} = 0.013$, $a_{14} = -0.01$, $a_{21} = 0.004$, $a_{22} = -0.022$, $a_{23} = -0.026$, $a_{24} = -0.024$, $a_{31} = 0.012$, $a_{32} = -0.013$, $a_{33} = -0.018$, $a_{34} = -0.015$, $a_{41} = 0.027$, $a_{42} = 0.002$, $a_{43} = -0.002$, $a_{44} = 0$, $b_{11} = 0$, $b_{12} = -0.003$, $b_{13} = -0.02$, $b_{14} = -0.023$, $b_{21} = 0.015$, $b_{22} = 0.012$, $b_{23} = -0.007$, $b_{24} = -0.008$, $b_{31} = 0.014$, $b_{32} = 0.011$, $b_{33} = -0.008$, $b_{34} = -0.01$, $b_{41} = 0.03$, $b_{42} = 0.028$, $b_{43} = 0.008$, $b_{44} = 0.007$

Table 9: Estimate of Ψ , estimated approximate standard error for $\hat{\Psi}$, and approximate 95% confidence interval for Ψ, applied to Table 3

Table	$\hat{\Psi}$	S.E.	C.I.
3	0.06	0.000036333	(0.0599 , 0.0601)

We can see from this measure that the degree of departure from MH for the data in Table 3 is estimated to be 6 percent of the maximum departure toward the lower-left-marginal inhomogeneity. All values in confidence interval for Ψ are positive. Therefore, Ψ is statistically significant and, the structure of MH model for educational level of man and woman departs toward the lower-left-marginal inhomogeneity.

Table 10: Measures of without each groups for 2011

2011	Ψ
First group removed	**0.0087**
Second group removed	0.0556
Third group removed	0.0788
Fourth group removed	0.0837

When the first group (primary school) is removed from the table, departure is the smallest (Table 10). This means that the departure from MH model is largest for the primary school graduates.

5. Discussions

The usual way analyzing the square contingency tables is to check the model fit. Additionally, measures of departure could be beneficial if the model does not hold for data. Because the measure of departure, exhibits the departure from the model and, can arise which level contribution on rejecting the hypothesis. From the results, we can say that, divorces in high school and university graduates both men and women have a more homogeneous structure whereas the primary and middle school graduates have a more heterogeneous structure. This means that the high school and university graduates got married to individuals who have the same level of education. When considering the elementary and middle school graduate's divorce distribution, it might be inferred that these individuals got married to people who have different levels of education. If we examine the overall frequency distribution of each of the three years, it can be seen that most divorces occur in individuals of primary school graduates.

6. References

[1] Agresti, A. 2010, *Analysis of ordinal categorical data* (Vol. 656). John Wiley & Sons.

[2] Stuart, A., 1955, A test for homogeneity of the marginal distributions in a two-way classification. *Biometrica* 42, 412-416

[3] Tahata, K., Iwashita, T., Tomizawa, S., 2008, Measure of departure from conditional marginal homogeneity for square contingency tables with ordered categories. *Statistics,* 42(5), 453-466

[4] Lawal, B, H, ,2008 On Fitting the Conditional Difference Asymmetry Models to Square Contingency Tables with Nominal Categories, *Quality & Quantity*, *42*(5), 605-612.

[5] Yamamoto, K., Ando, S., & Tomizawa, S, 2011. A Measure of Departure from Average Margınal Homogeneıty for Square Contıngency Tables wıth Ordered Categorıes. *REVSTAT–Statistical Journal*, *9*(2), 115-126.

[6] Tahata, K. , Kawasaki, K. , Tomizawa S., 201*)* Asymmetry Index on Marginal Homogeneity for Square Contingecy Tables with Ordered Categories. *Open Journal of Statistics*, Vol:2 Page:198-203

[7] Agresti, A., 2003 *Categorical data analysis* (Vol. 482). John Wiley & Sons.

[8] Turkish Statistical Institute (TÜİK) Marriage and Divorce Statistics 1991,2001,2011