Journal of Pediatric Sciences

A Review of Statistical Techniques for 2x2 and RxC Categorical Data Tables in SPSS

Cengiz BAL, Fikret ER, Harun SÖNMEZ

Journal of Pediatric Sciences; 1; e1

REVIEW ARTICLE

A Review of Statistical Techniques for 2x2 and RxC Categorical Data Tables In SPSS

Cengiz BAL[1], Fikret ER[2], Harun SÖNMEZ[3]

Abstract: *In this study, a review of statistical techniques for RxC categorical data tables is explained in detail. The emphasis is given to the association of techniques and their corresponding data considerations. Some suggestions to how to handle specific categorical data tables in SPSS and common mistakes in the interpretation of the SPSS outputs are shown.*

Keywords: *Categorical data, Chi square, Likelihood Ratio, Fisher's exact test, Yates, Mantel-Haenszel Test, Exact tests*
Received: 15/08/2009; Accepted: 13/10/2009

Introduction

Nowadays, with the rapid development in computer technology, the usage of statistical packages in many investigations is increased. Modern statistical softwares are designed to help the user to analyze his/her data efficiently and effectively. Ease of use of statistical software is important in order to increase the acceptance of the software in a wide range of scientific studies. At the same time, bringing a complex statistical procedure in to someones' fingertips by just clicking on some dialog boxes creates a different type of problem in which an untrained person (or a person without statistical knowledge to interpret the output) may think that he/she is the sole expert on the subject. The use of statistical techniques in any study should be carefully reviewed by an expert. The special importance should be given to "how the data is collected", "How it is summarized", and "Does the special assumptions for a given statistical technique satisfied by the data itself?". Biostatisticians suffer a lot from the problem of a non-expert person claiming a statistical fact based on their research findings without consulting a statistician about the whole of the project and the special considerations of the data.

In this study, categorical data tables are reviewed in terms of their analysis in SPSS using Chi-Square test; the emphasis is given to the outputs of SPSS, the identification of correct statistics and how to interpret the statistics obtained from the table.

Material and Methods

In this section the basics of categorical data tables and Chi-Square test are given. For this purpose, the main aim is to remind the reader the technique rather than to give techniqual details of the test.

Chi-Square Test

Chi-square statistic is used to assess the statistical significance of a finding. It is also used in goodness-of-fit tests. When a study needs to show that if there are statistically significant differences between the observed (or real) frequencies and the expected frequencies of two variables presented in a cross-tabulation or contingency table. In a contingency table, a table of frequencies classified according to two sets of values of categorical variables. It is called a contingency table because what you find in rows is contingent upon what you find in columns. One widely accepted interpretation of "no association" in a two-way contingency table is that row and column variables are independent [4, 5, 6, 8]. The classical test of hypothesis of independence is based on the chi-square statistic given in the following equation,

Corresponding Author: Cengiz BAL, PhD
Eskisehir Osmangazi University, Medical Faculty, Department of Biostatistics, 26480, Eskisehir, Turkey
Tel: + 90 222 2392979 – 4612
E-mail: cengiz@ogu.edu.tr
[1],Eskisehir Osmangazi University, Medical Faculty, Department of Biostatsitcis, 26480, Eskisehir, Turkey
[2,3],Anadolu University, Science Faculty, Department of Statistics, 26470, Eskisehir, Turkey

$$\chi_P^2 = \sum_{ij} \frac{\left(f_{ij} - E_{ij}\right)^2}{E_{ij}}$$

The main difficulty in using measures of association based on the chi-square statistic is that of finding a meaningful interpretation. The measures do not have simple probabilistic interpretations, and, despite the fact that chi-square is generally considered to be a good statistic for testing the hypothesis of independence, there is no consensus among statisticians that is also a good measure of association. In terms of squared difference between observed ands expected frequencies, the measures are useful for comparing several tables, but those whose ranges depend on the dimensions of the table are not really comparable across tables of different sizes [1, 3, 4, 6, 10].

Likelihood Ratio

It is also called the likelihood test or G test, is an alternative procedure to test the hypothesis of no association of columns and rows in nominal-level tabular data. It is supported by SPSS output and is based on maximum likelihood estimation. Though computed differently, likelihood ratio chi-square is interpreted the same way. For large samples, likelihood ratio chi-square will be close in results to Pearson chi-square.

$$\chi_{LR}^2 = -2\sum_{ij} f_{ij} \ln(E_{ij} / f_{ij})$$

Even for smaller samples, it rarely leads to different substantive results. SPSS will print likelihood ratio chi-square in the "Chi-Square Tests" table of output from the Analyze, Descriptives, Crosstabs menu selection. The equation that SPSS uses to calculate likelihood ratio chi-square statistic given as follows [4, 6, 12].

Fisher’s Exact Test

The Fisher exact test of significance may be used in place of the chi-square test in 2-by-2 tables, particularly for small samples. It tests the probability of getting a table as strong as the observed or stronger simply due to the chance of sampling, where "strong" is defined by the proportion of cases on the diagonal with the most cases [12]. Though usually employed as a one-tailed test, it may be computed as a two-tailed test as well. The Fisher exact test is sometimes known as the Fisher-Irwin test as it was developed at the same time by Fisher, Irwin, and Yates in the 1930's. SPSS automatically computes Fisher's exact test in addition to chi-square for 2x2 tables when the table has a cell with an expected frequency of less than 5. Consider the following observed 2x2 table;

n_1	n_2	**$n_1 + n_2$**
n_3	n_4	**$n_3 + n_4$**
$n_1 + n_3$	**$n_2 + n_4$**	**N**

Conditional on the observed marginal totals, the values of the four cell counts can be expressed as the observed count of the first cell n_1 only. Under the hypothesis of independence, the count of the first cell N_1 follows a hypergeometric distribution with the probability of $N_1 = n_1$ given by

$$\Pr ob(N_1 = n_1) = \frac{(n_1 + n_2)!(n_3 + n_4)!(n_1 + n_3)!(n_2 + n_4)!}{N!n_1!n_2!n_3!n_4!}$$

where N1 ranges from max(0, n1 ... n4) to min(n1+n2, n1+n3) and N = n1+n2+n3+n4.

The exact one-tailed significance level p_1 is defined as,

$$p_1 = \begin{cases} \Pr ob(N_1 \geq n_1) & \text{if } n_1 > E(N_1) \\ \Pr ob(N_1 \leq n_1) & \text{if } n_1 \leq E(N_1) \end{cases}$$

where

$$E(N_1) = (n_1 + n_2)(n_1 + n_3) / N.$$

The exact two-tailed significance level p_2 is defined as the sum of the one-tailed significance level p_1 and the probabilities of all points in the other side of the sample space of N_1 which are not greater than the probability of $N_1 = n_1$ [4, 6, 7].

Yates Continuity Correction for 2x2 Tables

Yates’s correction (Yates, 1934, 1984) is used as an approximation in the analysis of 2×1 and 2×2 contingency tables. A 2×2 contingency table shows the frequencies of occurrence of all combinations of the levels of two dichotomous variables, in a sample of size N. A research question of interest is often whether the variables summarized in a contingency table are independent of each other. The test to determine if this is so depends on which, if any, of the margins are fixed, either by design or for the purposes of the analysis. The equation that SPSS uses is given as follows.

$$\chi_C^2 = \begin{cases} \dfrac{W\left(\left|f_{11} f_{22} - f_{12} f_{21}\right| - 0.5W\right)^2}{r_1 r_2 c_1 c_2} & if\ \left|f_{11} f_{22} - f_{12} f_{21}\right| > 0{,}5W \\ 0 & otherwise \end{cases}$$

The degrees of freedom is 1 [4, 6, 11].

Mantel-Haenszel Test of Linear Association

It is also called the Mantel-Haenszel test for linear association or linear by linear association chi-square, unlike ordinary and likelihood ratio chi-square, is an ordinal measure of significance. It is preferred when testing the significance of linear relationship between two ordinal variables because it is more powerful than Pearson chi-square (more likely to establish linear association). Mantel-Haenzel chi-square is not appropriate for nominal variables. If found significant, the interpretation is that increases in one variable are associated with increases (or decreases for negative relationships) in the other greater than would be expected by chance of random sampling. Like other chi-square statistics, M-H chi-square should not be used with tables with small cell counts. The equation that SPSS uses is given as follows.

$$\chi^2_{MH} = (W-1)r^2$$

where r is the Pearson correlation coefficient. The degree of freedom is 1 [4, 5, 6].

Exact tests

The goal in Exact Tests is to enable a researcher to make reliable inferences when the data is small, sparse, heavily tied, or unbalanced and the validity of the corresponding large sample theory is in doubt. This is achieved by computing exact *p* values for a very wide class of hypothesis tests, including one-, two-, and K-sample tests, tests for unordered and ordered categorical data, and tests for measures of association. The statistical methodology underlying these exact tests is well established in the statistical literature and may be regarded as a natural generalization of Fisher's exact test for the single 2X2 contingency table. The real challenge has been to make this methodology operational through software development. For small data sets, the algorithms ensure quick computation of exact *p* values. If a data set is too large for the exact algorithms, Monte Carlo algorithms are substituted in their place in order to estimate the exact p values to any desired level of accuracy [7].

The Exact Tests option provides two new methods for calculating significance levels for the statistics available through the Crosstabs and Nonparametric Tests procedures in SPSS. These new methods, the exact and Monte Carlo methods, provide a powerful means for obtaining accurate results when your data set is small, your tables are sparse or unbalanced, the data are not normally distributed, or the data fail to meet any of the underlying assumptions necessary for reliable results using the standard asymptotic method [7].

The Exact Method

By default, SPSS Statistics calculates significance levels for the statistics in the Crosstabs and Nonparametric Tests procedures using the asymptotic method. This means that *p* values are estimated based on the assumption that the data, given a sufficiently large sample size, conform to a particular distribution. However, when the data set is small, sparse, contains many ties, is unbalanced, or is poorly distributed, the asymptotic method may fail to produce reliable results. In these situations, it is preferable to calculate a significance level based on the exact distribution of the test statistic. This enables a researcher to obtain an accurate *p* value without relying on assumptions that may not be met by your data [7].

EXAMPLES

In this study the data comes from Turkey's Statistical Year Book 2008, so the responsibility of the correctness of the data belongs to State Statistical Institute. Special importance is given to some well-known diseases that one can encounter in daily life [9].

RXC TABLES

In the first example 3 different diseases are classified into contingency table where the columns of the table is the year variable. The data is shown in Table 1.

Table 1: The number of cases for selected infection diseases obligatory to be reported in Turkey by year.

	2004	**2005**	**2006**	**Total**
Tetanus	22	19	10	**51**
Meningitis	611	216	183	**1010**
Paratyphoid fever	429	263	155	**847**
Total	**1062**	**498**	**348**	**1908**

Pearson chi-square statistic should be computed in r x c tables. Pearson chi-square statistics is computed when a table does not contain missing cells for rows or columns, and the contingency table should not contain a cell in which the expected frequency being less than 5. The null hypothesis for the Person chi-square test is written as the row and column variables are independent of each other.

Explanation of "a" below SPSS output Table 2 (see Appendix) indicates that the minimum expected count is 9.30, and non of the cells have expected counts less than 5.

The computed chi-square for SPSS output Table 2 (see Appendix) is 28.545 and its associated probability (p value) is shown as 0.000, but this 0.000 do not indicate that it is exactly equals to zero. It means that the number is to small to be shown in the table. The important thing is that this probability value is less than the significance level that can be used in an investigation such as 0.05. Since in this example the probability value is less than 0.05/0.01 significance level the hypothesis of independence is rejected. Distribution of reported AIDS cases and carriers by age groups and sex in Turkey for children under 14 in 1985-2006. [1]

In the second example the contingency table contains the number AIDS cases and carriers by age group and sex for children under 14 years in Turkey. The data is shown in Table 2.

Table 2: Distribution of reported AIDS cases and carriers by age groups and sex in Turkey for children under 14 between 1985-2006.

Sex/Ages	0 year	1-4 years	5-9 years	10-12 years	13-14 years	Total
Females	4	12	8	2	1	**25**
Males	13	6	3	3	1	**26**
Total	**17**	**18**	**11**	**5**	**2**	**51**

According to former statements, explanation of “a” below SPSS Output Table 4a (see Appendix) indicates that the minimum expected count is 0.98, and 4 cells have expected counts less than 5. In this situation the exact Pearson chi-square statistic should be calculated and be used in conclusion statements.

The computed chi-square statistics for SPSS output Table 4a and 4b (see Appendix) is 9.222. In both of the tables the calculated Pearson chi-square statistic is the same but the extra explanations below the tables give a warning. According to warning “a” the usage of Pearson chi-square statistic is not correct. In view of the explanations under Table 4a and 4b Monte Carlo significance value should be used and it is equal to 0.046 (which is less than a presumable alpha level 0.05). If this statistic is used then null hypothesis is rejected but in the other case it is easy to see that the researcher should accept the null hypothesis. Clearly the correct technique is extremely important here since a change of decision might occur.

2X2 TABLES

In the third example the contingency table contains the number AIDS cases and cariers by years in Turkey. The data is shown in Table 3.

Table 3: Distribution of reported AIDS cases and carriers in Turkey for 2005-2006.

	Year 2005	Year 2006	Total
Cases	37	35	72
Carriers	295	255	550
Total	332	290	622

The table 3 is arranged as 2x2 contingency table. Explanation of “b” below SPSS output table 6 (see Appendix) indicates that the minimum expected count is 33.57, none of the cells have expected count less than 5.

The computed chi-square statistics for SPSS output Table 6 (see Appendix) is 0.129 and its associated probability (p value) is calculated as 0.719. Since the probability value is greater than 0.05/0.01 significance level the hypothesis of independence is accepted.

In the fourth example the contingency table contains the number of cases of selected infection diseases obligatory to be reported, Tetanus and Whooping cough in Turkey 2005-2006. The data is shown in Table 4.

Table 4: Number of cases of selected infection diseases obligatory to be reported, Tetanus and Whooping cough, in Turkey 2005-2006.

	Year 2005	Year 2006	Total
Tetanus	19	10	**29**
Whooping cough	72	57	**129**
Total	**91**	**67**	**158**

The table 4 is arranged as 2x2 contingency table. Explanation of “b” below SPSS output table 8 (see Appendix) indicates that the minimum expected count is 12.30, non of the cells have expected count less than 5. For this reason the continuity correction value and associated probability should be used for hypothesis procedure (If any cells’ value is between 5 and 25, continuity correction value and associated probability may be used).

The computed continuity correction value for SPSS output Table 8 (see Appendix) is 0.559 and its associated probability (p value) is calculated as 0.445. Since the probability value is greater than 0.05/0.01 significance level the hypothesis of independence is accepted.

In the fifth example the contingency table contains distribution of reported Tetanus and Diphtheria in Turkey 2002-2003. The data is shown in Table 5.

Table 5: Distribution of reported Tetanus and Diphtheria in Turkey 2002-2003.

	Year 2002	Year 2003	Total
Diphtheria	2	1	**3**
Tetanus	16	17	**33**
Total	**18**	**18**	**36**

The table 5 is arranged as 2x2 contingency table. Explanation of “b” below SPSS output table 10 (see Appendix) indicates that the minimum expected count is 1.50, two cells have expected count less than 5. For this reason Fisher’s exact test probability (exact significance (2-sided) or (1-sided)) should be used for hypothesis procedure (If any cells' value is below 5, Fisher's exact probability may be used).

The computed Fisher's exact test probability for SPSS output Table 10 (see Appendix)is 1.00. Since the probability value is greater than 0.05/0.01 significance level the hypothesis of independence is accepted.

RESULTS

Nowadays, statistical softwares used in scientific studies are very fast. The computer technology allows the software developers to develop better algorithms than before. At the same time there is an increase on the number of studies involve two or more scientific area working together.

The increase in the number of statistical software can be thought as a good thing for a beginning but the reality becomes a little bit complicated. Especially in interdisciplinary studies the problem is more obvious. A person without a theoretical background on a given subject can easily solve the problem via a simple software, but when an expert looks at the problem he/she may see that most of the assumptions are not satisfied for the technique used in the analysis by the data or study design. Even though the investigator obtains the statistical output from the computer and interpret it correctly, if one does not talk about the procedure in which the investigator followed during the study then the output may become redundant. Especially if the output of the study is used in a health care problem then the mistakes caused by insufficient statistical considerations of the analysis design become more dangerous and may end up with some critical problems to the public health. Statisticians, especially biostatisticians, always draw the point that the statistical techniques should be carefully used in any stage of a study design regardless of the techniques ease of calculation.

The chi-square tests that are reviewed in this study are frequently used by researchers. In many studies, ignoring the general structure of a contingency table, researchers calculate and interpret Pearson chi-square statistic. On some cases the researcher does not look at the number of observations in each cell and even with zero frequency in one or two cells it has been seen that the chi-square statistic is calculated and interpreted. When this happens it is very difficult to say that the findings of the study are well founded. In order to decrease these kind of problems it is necessary to include a biostatistician to comment on the general statistical aspects of the study.

In this study, real life examples from Turkish health care data for 2008 are designed and shown as r x c and 2 x 2 contingency tables and appropriate test shown via SPSS. Then the output obtained from SPSS is investigated in detail.

The purpose of the study is not to teach chi-square tests. Rather the main aim is to remind the reader that most of the time Pearson chi-square test is taken granted as the best simple technique which can be applied in to any contingency table. In fact careful examination of the variables involved in the study and the number of observations in each cell can lead an investigator in to another test, and even in to some different conclusion than Pearson chi-square test.

APPENDIX

SPSS Output Table 1: Crosstabulation For Disease And Year

Diseases * Years Crosstabulation

Count

		Years			Total
		2004	2005	2006	
Diseases	Tetanus	22	19	10	51
	Menengitis	611	216	183	1010
	Paratyphoid fever	429	263	155	847
Total		1062	498	348	1908

SPSS Output Table 2: Chi-Square Test Results For SPSS Output Table 1.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Monte Carlo Sig. (2-sided) Sig.	Monte Carlo Sig. (2-sided) 99% Confidence Interval Lower Bound	Monte Carlo Sig. (2-sided) 99% Confidence Interval Upper Bound	Monte Carlo Sig. (1-sided) Sig.	Monte Carlo Sig. (1-sided) 99% Confidence Interval Lower Bound	Monte Carlo Sig. (1-sided) 99% Confidence Interval Upper Bound
Pearson Chi-Square	28,545[a]	4	,000	,000[b]	,000	,000			
Likelihood Ratio	28,492	4	,000	,000[b]	,000	,000			
Fisher's Exact Test	28,658			,000[b]	,000	,000			
Linear-by-Linear Association	3,795[c]	1	,051	,057[b]	,051	,063	,029[b]	,025	,034
N of Valid Cases	1908								

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 9,30.

b. Based on 10000 sampled tables with starting seed 92208573.

c. The standardized statistic is 1,948.

SPSS Output Table 3: Crosstabulation For Sex And Age Groups.

Sex * Age_Groups Crosstabulation

Count

		Age_Groups					Total
		0	1-4 years	5-9 years	10-12 years	13-14 years	
Sex	Female	4	12	8	2	1	27
	Male	13	6	3	3	1	26
Total		17	18	11	5	2	53

SPSS Output Table 4a: Chi-Square Test Results For SPSS Output Table 3.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9,222[a]	4	,056
Likelihood Ratio	9,596	4	,048
Linear-by-Linear Association	2,136	1	,144
N of Valid Cases	53		

a. 4 cells (40,0%) have expected count less than 5. The minimum expected count is ,98.

SPSS Output Table 4b: Chi-Square Test Results For SPSS Output Table 3.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Monte Carlo Sig. (2-sided) Sig.	Monte Carlo Sig. (2-sided) 99% Confidence Interval Lower Bound	Monte Carlo Sig. (2-sided) 99% Confidence Interval Upper Bound	Monte Carlo Sig. (1-sided) Sig.	Monte Carlo Sig. (1-sided) 99% Confidence Interval Lower Bound	Monte Carlo Sig. (1-sided) 99% Confidence Interval Upper Bound
Pearson Chi-Square	9,222[a]	4	,056	,046[b]	,041	,051			
Likelihood Ratio	9,596	4	,048	,075[b]	,068	,081			
Fisher's Exact Test	9,354			,034[b]	,029	,038			
Linear-by-Linear Association	2,136[c]	1	,144	,177[b]	,167	,186	,089[b]	,082	,096
N of Valid Cases	53								

a. 4 cells (40,0%) have expected count less than 5. The minimum expected count is ,98.

b. Based on 10000 sampled tables with starting seed 2000000.

c. The standardized statistic is -1,461.

SPSS Output Table 5: Crosstabulation for AIDS (cases and carriers) and year.

AIDS * Years Crosstabulation

Count

		Years 2005	Years 2006	Total
AIDS	Cases	37	35	72
	Carriers	295	255	550
Total		332	290	622

SPSS Output Table 6: Chi-Square test results for SPSS Output Table 5

Chi-Square Tests[d]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,129[b]	1	,719	,802	,407	
Continuity Correction[a]	,055	1	,815			
Likelihood Ratio	,129	1	,719	,802	,407	
Fisher's Exact Test				,802	,407	
Linear-by-Linear Association	,129[c]	1	,719	,802	,407	,093
N of Valid Cases	622					

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 33,57.

c. The standardized statistic is -,359.

d. For 2x2 crosstabulation, exact results are provided instead of Monte Carlo results.

SPSS Output Table 7: Crosstabulation for Diseases and Year

Diseases * Years Crosstabulation

Count

		Years 2005	Years 2006	Total
Diseases	Tetanus	19	10	29
	Whooping cough	72	57	129
Total		91	67	158

SPSS Output Table 8: Chi-Square test results for SPSS Output Table

Chi-Square Tests[d]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,913[b]	1	,339	,408	,229	
Continuity Correction[a]	,559	1	,455			
Likelihood Ratio	,928	1	,335	,408	,229	
Fisher's Exact Test				,408	,229	
Linear-by-Linear Association	,907[c]	1	,341	,408	,229	,107
N of Valid Cases	158					

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 12,30.

c. The standardized statistic is ,952.

d. For 2x2 crosstabulation, exact results are provided instead of Monte Carlo results.

SPSS Output Table 9: Cross-tabulation for diseases and year.

Diseases * Years Crosstabulation

Count

		Years 2002	Years 2003	Total
Diseases	Diphtheria	2	1	3
	Tetanus	16	17	33
Total		18	18	36

Output table 10: Chi-Square test results for SPSS Output Table 9

Chi-Square Tests[d]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,364[b]	1	,546	1,000	,500	
Continuity Correction[a]	,000	1	1,000			
Likelihood Ratio	,370	1	,543	1,000	,500	
Fisher's Exact Test				1,000	,500	
Linear-by-Linear Association	,354[c]	1	,552	1,000	,500	,386
N of Valid Cases	36					

a. Computed only for a 2x2 table

b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 1,50.

c. The standardized statistic is ,595.

d. For 2x2 crosstabulation, exact results are provided instead of Monte Carlo results.

REFERENCES

1. Agresti, Alan (1996). "Introduction to categorical data analysis" NY: John Wiley and Sons

2. Garson, D. "Statnotes: Topics in Multivariate Analysis", September, 2009 "http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm".

3. Liebetrau, A.M. (1983), "Measures of Association", A sage university paper, Sage Publications, 1983.

4. Özdamar, K (2003). "SPSS ile Biyoistatistik", Kaan Kitabevi, Eskişehir, Turkey (In Turkish).

5. Richardson, J.T.E. (1994) The analysis of 2 × 1 and 2 × 2 contingency tables: an historical review. Statistical Methods in Medical Research, 3, 107–133.

6. Spss 15.0 Algorithms, SPSS Inc., USA, 2006. (Downloaded: September 2009, http://support.spss.com.)

7. Metha, C. R., Patel, N. R.: Exact Tests (SPSS user guide), SPSS Inc., USA. (Downloaded: September 2009, http://support.spss.com.)

8. Siegel, S. and Castellan, N.J. (1988) Nonparametric Statistics for the Behavioural Sciences. 2nd Edition, New York: McGraw–Hill.

9. Turkish Statistical Institute, "Turkey's Statistical Year Book 2008", Publication Number: 3248, Ankara, Turkey.

10. Vogt, W.P. (1993) "Dictionary of Statistics and Methodology", Sage Publications.

11. Yates, F. (1934) Contingency tables involving small numbers and the χ^2 test. Journal of the Royal Statistical Society, Suppl.1, 217–235.

12. Yates, F. (1984) Tests of significance for 2 × 2 contingency tables. Journal of the Royal Statistical Society Series A, 147, 426–463.