Chapter 501

Contingency Tables (Crosstabs / Chi-Square Test)

Introduction

This procedure produces tables of counts and percentages for the joint distribution of two categorical variables. Such tables are known as *contingency*, *cross-tabulation*, or *crosstab* tables. When a breakdown of more than two variables is desired, you can specify up to eight grouping (break) variables in addition to the two table variables. A separate table is generated for each unique set of values of these grouping variables. The data can also be entered directly as a two-way table for analysis.

This procedure serves as a summary reporting tool and is often used to analyze survey data. It calculates most of the popular contingency-table statistics and tests such as chi-square, Fisher's exact, and McNemar's tests, as well as the Cochran-Armitage test for trend in proportions and the Kappa and weighted Kappa tests for inter-rater agreement. It also calculates pairwise multiple comparisons of proportions as well as Dunnett-type multiple comparisons to a control.

This procedure also produces a broad set of association and correlation statistics for contingency tables: Phi, Cramer's V, Pearson's Contingency Coefficient, Tschuprow's T, Lamba, Kendall's Tau, and Gamma.

Types of Categorical Variables

Note that we will refer to two types of categorical variables: *Table* variables and *Grouping* variables. The values of the *Table* variables are used to define the rows and columns of a single contingency table. Two *Table* variables are used for each table, one variable defining the rows of the table and the other defining the columns. *Grouping* variables are used to split a data into subgroups. A separate table is generated for each unique set of values of the *Grouping* variables.

Note that if you only want to use one *Table* variable, you should use the *Frequency Table* procedure.

Technical Details

For the technical details that follow, we assume a contingency table of counts with R rows and C columns as in the table below. Let O_{ij} be the observed count for the i^{th} row (i = 1 to R) and j^{th} column (j = 1 to C). Let the row

		Column Variable					
		Column 1	$\cdots$	**Column *j***	$\cdots$	**Column *C***	**Total**
Row Variable	**Row 1**	O_{11}	$\cdots$	O_{1j}	$\cdots$	O_{1C}	$n_{1\cdot}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	⋰	$\vdots$	$\vdots$
	Row *i*	O_{i1}	$\cdots$	O_{ij}	$\cdots$	O_{iC}	$n_{i\cdot}$
	$\vdots$	$\vdots$	⋰	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	Row *R*	O_{R1}	$\cdots$	O_{Rj}	$\cdots$	O_{RC}	$n_{R\cdot}$
	Total	$n_{\cdot 1}$	$\cdots$	$n_{\cdot j}$	$\cdots$	$n_{\cdot C}$	1

and column marginal totals be designated as $n_{i\cdot}$ and $n_{\cdot j}$, respectively, where

$$n_{i\cdot} = \sum_j O_{ij}$$

$$n_{\cdot j} = \sum_i O_{ij}$$

Let the total number of counts in the table be N, where

$$\begin{aligned} N &= \sum_i \sum_j O_{ij} \\ &= \sum_i n_{i\cdot} \\ &= \sum_j n_{\cdot j} \end{aligned}$$

The table of associated proportions can then be written as

		Column Variable					
		Column 1	$\cdots$	**Column *j***	$\cdots$	**Column *C***	**Total**
Row Variable	**Row 1**	p_{11}	$\cdots$	p_{1j}	$\cdots$	p_{1C}	$p_{1\cdot}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	⋰	$\vdots$	$\vdots$
	Row *i*	p_{i1}	$\cdots$	p_{ij}	$\cdots$	p_{iC}	$p_{i\cdot}$
	$\vdots$	$\vdots$	⋰	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	Row *R*	p_{R1}	$\cdots$	p_{Rj}	$\cdots$	p_{RC}	$p_{R\cdot}$
	Total	$p_{\cdot 1}$	$\cdots$	$p_{\cdot j}$	$\cdots$	$p_{\cdot C}$	1

where

$$p_{ij} = \frac{O_{ij}}{N}$$

$$p_{i\cdot} = \frac{n_{i\cdot}}{N}$$

$$p_{\cdot j} = \frac{n_{\cdot j}}{N}$$

Finally, designate the expected counts and expected proportions for the i^{th} row and j^{th} column as E_{ij} and Pe_{ij}, respectively, where

$$E_{ij} = \frac{n_{i\cdot}n_{\cdot j}}{N}$$

$$P_{e_{ij}} = \frac{E_{ij}}{N} = p_{i\cdot}p_{\cdot j}$$

In the sections that follow we will describe the various tests and statistics calculated by this procedure using the preceding notation.

Table Statistics

This section presents various statistics that can be output for each individual cell. These are useful for studying the independence between rows and columns. The statistics for the i^{th} row and j^{th} column are as follows.

Count

The cell count, O_{ij}, is the number of observations for the cell.

Row Percentage

The percentage for column j within row i, $p_{j|i}$, is calculated as

$$p_{j|i} = \frac{O_{ij}}{n_{i\cdot}}$$

Column Percentage

The percentage for row i within column j, $p_{i|j}$, is calculated as

$$p_{i|j} = \frac{O_{ij}}{n_{\cdot j}}$$

Table Percentage

The overall percentage for the cell, p_{ij}, is calculated as

$$p_{ij} = \frac{O_{ij}}{N}$$

Expected Counts Assuming Independence

The expected count, E_{ij}, is the count that would be obtained if the hypothesis of row-column independence were true. It is calculated as

$$E_{ij} = \frac{n_{i\cdot}n_{\cdot j}}{N}$$

Chi-Square Contribution

The chi-square contribution, CS_{ij}, measures the amount that a cell contributes to the overall chi-square statistic for the table. This and the next two items let you determine which cells impact the chi-square statistic the most.

$$CS_{ij} = \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

Deviation from Independence

The deviation statistic, D_{ij}, measures how much the observed count differs from the expected count.

$$D_{ij} = O_{ij} - E_{ij}$$

Std. Residual

The standardized residual, SR_{ij}, is equal to the deviation divided by the square root of the expected value:

$$SR_{ij} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}}$$

Tests for Row-Column Independence

Pearson's Chi-Square Test

Pearson's chi-square statistic is used to test independence between the row and column variables. Independence means that knowing the value of the row variable does not change the probabilities of the column variable (and vice versa). Another way of looking at independence is to say that the row percentages (or column percentages) remain constant from row to row (or column to column).

This test requires large sample sizes to be accurate. An often-quoted rule of thumb regarding sample size is that *none of the expected cell values should be less than five*. Although some users ignore the sample size requirement, you should also be very skeptical of the test if you have cells in your table with zero counts. For 2 × 2 tables, consider using *Yates' Continuity Correction* or *Fisher's Exact Test* for small samples.

Pearson's chi-square test statistic follows an asymptotic chi-square distribution with $(R-1)(C-1)$ degrees of freedom when the row and column variables are independent. It is calculated as

$$\chi_P^2 = \sum_i \sum_j \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}.$$

Yates' Continuity Corrected Chi-Square Test (2 × 2 Tables)

Yates' Continuity Corrected Chi-Square Test (or just Yates' Continuity Correction) is similar to Pearson's chi-square test but is adjusted for the continuity of the chi-square distribution. This test is particularly useful when you have small sample sizes. This test is only calculated for 2 × 2 tables.

Yates' continuity corrected test statistic follows an asymptotic chi-square distribution with $(R-1)(C-1)$ degrees of freedom when the row and column variables are independent. It is calculated as

$$\chi_Y^2 = \sum_i \sum_j \frac{\left(max\left(0, \left|O_{ij} - E_{ij}\right| - 0.5\right)\right)^2}{E_{ij}}.$$

Likelihood Ratio Test

This test makes use of the fact that under the null hypothesis of independence, the likelihood ratio statistic follows an asymptotic chi-square distribution.

The likelihood ratio test statistic follows an asymptotic chi-square distribution with $(R-1)(C-1)$ degrees of freedom when the row and column variables are independent. It is calculated as

$$\chi_{LR}^2 = 2\sum_i \sum_j O_{ij} \ln\left(\frac{O_{ij}}{E_{ij}}\right).$$

Fisher's Exact Test (2 × 2 Tables)

This test was designed to test the hypothesis that the two column percentages in a 2 × 2 table are equal. It is especially useful when sample sizes are small (even zero in some cells) and the chi-square test is not appropriate.

Using the hypergeometric distribution with fixed row and column totals, this test computes probabilities of all possible tables with the observed row and column totals. This test is often used when sample sizes are small, but it is appropriate for all sample sizes because Fisher's exact test does not depend on any large-sample asymptotic distribution assumptions. This test is only calculated for 2 × 2 tables.

If we assume that P_H is the hypergeometric probability of any table with the observed row and column marginal totals, then Fisher's Exact Test probabilities are calculated by summing over defined sets of tables depending on the hypothesis being tested (one-sided or two-sided).

Define the difference between conditional column proportions for row 1 from the observed table as D_O, with

$$D_O = p_{1|1} - p_{1|2}$$

and the difference between conditional column proportions for row 1 from other possible tables with the observed row and column marginal totals as D, with

$$D = p_{1|1} - p_{1|2}$$

The two-sided Fisher's Exact Test P-value is calculated as

$$P - Value_{Two-Sided} = \sum_{Tables\ where\ |D| \geq |D_O|} P_H$$

The lower one-sided Fisher's Exact Test P-value is calculated as

$$P - Value_{Lower} = \sum_{Tables\ where\ D \leq D_O} P_H$$

The upper one-sided Fisher's Exact Test P-value is calculated as

$$P - Value_{Upper} = \sum_{Tables\ where\ D \geq D_O} P_H$$

Tests for Trend in Proportions (2 × *k* Tables)

When one variable is ordinal (e.g., "Low, Medium, High" or "1, 2, 3, 4, 5") and the other has exactly two levels (e.g., "success", "failure"), you can test the hypothesis that there is a linear trend in the proportion of successes (i.e., that the true proportion of successes increases (or decreases) across the levels of the ordinal variable). Three tests for linear trend in proportions are available in **NCSS**: the Cochran-Armitage Test, the Cochran-Armitage Test with Continuity Correction, and the Armitage Rank Correlation Test. Of these, the Cochran-Armitage Test is the most widely used.

Cochran-Armitage Test

The Cochran-Armitage test is described in Cochran (1954) and Armitage (1955). Though the formulas that follow appear different from those presented in the articles, the results are equivalent.

Suppose we have *k* independent binomial variates, y_i, with response probabilities, p_i, based on samples of size n_i at covariate (or dose) levels, x_i, for i = 1, 2, …, *k*, where $x_1 < x_2 < ... < x_k$. The scores x_i, come from the row (or column) names of the ordinal variable. When the names are numeric (e.g., "1 2 3 4 etc.") then the actual numeric values are used for the scores, allowing the user to input unequally spaced score values. When the names are not numeric, even though they may represent an ordinal scale (e.g., "Low, Medium, High"), then the scores are assigned automatically as evenly spaced integers from 1 to *k*.

Define the following:

$$N = \sum_{i=1}^{k} n_i$$

$$\overline{p} = \frac{1}{N}\sum_{i=1}^{k} y_i$$

$$\overline{q} = 1 - \overline{p}$$

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{k} n_i x_i$$

If we assume that the probability of response follows a linear trend on the logistic scale, then

$$p_i = \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)}.$$

The Cochran-Armitage test can be used to test the following hypotheses:

One-Sided (Increasing Trend)	$H_0: p_1 = p_2 = ... = p_k$ vs. $H_1: p_1 < p_2 < ... < p_k$
One-Sided (Decreasing Trend)	$H_0: p_1 = p_2 = ... = p_k$ vs. $H_1: p_1 > p_2 > ... > p_k$
Two-Sided	$H_0: p_1 = p_2 = ... = p_k$ vs. $H_1: p_1 < p_2 < ... < p_k$ or $p_1 > p_2 > ... > p_k$

Nam (1987) presents the following asymptotic test statistic for detecting a linear trend in proportions

$$z = \frac{\sum_{i=1}^{k} y_i (x_i - \overline{x})}{\sqrt{\overline{pq}\left[\sum_{i=1}^{k} n_i (x_i - \overline{x})^2\right]}}.$$

A one-sided test rejects H_0 in favor of an <u>increasing trend</u> if $z \geq z_{1-\alpha}$, where $z_{1-\alpha}$ is the value that leaves 1 – α in the upper tail of the standard normal distribution. A one-sided test rejects H_0 in favor of a <u>decreasing trend</u> if $z \leq z_{\alpha}$, where z_{α} is the value that leaves α in the lower tail of the standard normal distribution. A two-sided test rejects H_0 in favor of either an increasing or decreasing trend if $|z| \geq z_{1-\alpha/2}$.

Cochran-Armitage Test with Continuity Correction

The Cochran-Armitage test with continuity correction is nearly the same as the uncorrected Cochran-Armitage test described earlier. In the continuity corrected test, a small continuity correction factor, $\Delta/2$, is added or subtracted from the numerator, depending on the direction of the test. If the scores, x_i, are equally spaced then

$$\Delta = x_{i+1} - x_i \text{ for all } i < k$$

or the interval between adjacent scores. **NCSS** computes Δ for unequally spaced scores as

$$\Delta = \frac{1}{k-1}\sum_{i=1}^{k-1}\left(x_{i+1} - x_i\right).$$

For the case of unequally spaced covariates, Nam (1987) states, "For unequally spaced doses, no constant correction is adequate for all outcomes." Therefore, we caution against the use of the continuity-corrected test statistic in the case of unequally spaced covariates.

Using the same notation as that described for the Cochran-Armitage test, Nam (1987) presents the following continuity corrected asymptotic test statistic for detecting an increasing linear trend in proportions

$$z_{c.c.U} = \frac{\sum_{i=1}^{k} y_i\left(x_i - \bar{x}\right) - \frac{\Delta}{2}}{\sqrt{\overline{pq}\left[\sum_{i=1}^{k} n_i\left(x_i - \bar{x}\right)^2\right]}}.$$

A one-sided test rejects H_0 in favor of an <u>increasing trend</u> if $z_{c.c.U} \geq z_{1-\alpha}$, where $z_{1-\alpha}$ is the value that leaves 1 – α in the upper tail of the standard normal distribution.

The continuity-corrected test statistic for a decreasing trend is the same as that for an increasing trend, except that $\Delta/2$ is added in the numerator instead of subtracted

$$z_{c.c.L} = \frac{\sum_{i=1}^{k} y_i\left(x_i - \bar{x}\right) + \frac{\Delta}{2}}{\sqrt{\overline{pq}\left[\sum_{i=1}^{k} n_i\left(x_i - \bar{x}\right)^2\right]}}.$$

A one-sided test rejects H_0 in favor of a <u>decreasing trend</u> if $z_{c.c.L} \leq z_{\alpha}$, where z_{α} is the value that leaves α in the lower tail of the standard normal distribution.

A two-sided test rejects H_0 in favor of either an increasing or decreasing trend if $z_{c.c.U} \geq z_{1-\alpha/2}$ or if $z_{c.c.L} \leq z_{\alpha/2}$.

Armitage Rank Correlation Test

The Armitage Rank Correlation test is described in section 4 of Armitage (1955) (the test is referred to as *Kendall's Rank Correlation Test* in the paper). The statistic, S, is standardized to a normal z-value by dividing by the estimated standard error of S (which we label $\sqrt{V}$ below). This z-value can be tested using the standard-normal distribution.

When there are two columns and we want to test for the presence of a trend in proportions down the rows, the calculations for this test are as follows:

$$z = \frac{S}{\sqrt{V}}$$

where

$$S = A - B$$

$$V = \frac{n_{\cdot 1} n_{\cdot 2}\left(N^3 - \sum_{i=1}^{R} n_{i\cdot}^3\right)}{3N(N-1)}$$

with

$$A = \sum_{j=1}^{R-1} O_{j2} \sum_{i=j+1}^{R} O_{i1}$$

$$B = \sum_{j=1}^{R-1} O_{j1} \sum_{i=j+1}^{R} O_{i2}$$

A one-sided test rejects H_0 in favor of an increasing trend if $z \geq z_{1-\alpha}$, where $z_{1-\alpha}$ is the value that leaves $1-\alpha$ in the upper tail of the standard normal distribution. A one-sided test rejects H_0 in favor of a decreasing trend if $z \leq z_{\alpha}$, where z_{α} is the value that leaves α in the lower tail of the standard normal distribution. A two-sided test rejects H_0 in favor of either an increasing or decreasing trend if $|z| \geq z_{1-\alpha/2}$.

McNemar Test (*k* × *k* Tables)

The McNemar test was first used to compare two proportions that are based on matched samples. Matched samples occur when individuals (or matched pairs) are given two different treatments, asked two different questions, or measured in the same way at two different points in time. Match pairs can be obtained by matching individuals on several other variables, by selecting two people from the same family (especially twins), or by dividing a piece of material in half.

The McNemar test has been extended so that the measured variable can have more than two possible outcomes. It is then called the *McNemar test of symmetry*. It tests for symmetry around the diagonal of the table. The diagonal elements of the table are ignored. The test is computed for square $k \times k$ tables only.

The McNemar test statistic follows an asymptotic chi-square distribution with $R(R-1)/2$ degrees of freedom. It is calculated as

$$\chi_M^2 = \frac{1}{2}\sum_i \sum_j \frac{\left(O_{ij} - O_{ji}\right)^2}{\left(O_{ij} + O_{ji}\right)}.$$

Kappa and Weighted Kappa Tests for Inter-Rater Agreement (*k* × *k* Tables)

Kappa is a measure of association (correlation or reliability) between two measurements on the same individual when the measurements are categorical. It tests if the counts along the diagonal are significantly large. Because Kappa is used when the same variable is measured twice, it is only appropriate for square tables where the row and column categories are the same. Kappa is often used to study the agreement of two raters such as judges or doctors, where each rater classifies each individual into one of *k* categories.

Rules-of-thumb for kappa: values less than 0.40 indicate low association; values between 0.40 and 0.75 indicate medium association; and values greater than 0.75 indicate high association between the two raters.

Kappa and weighted kappa are only output for square *k* × *k* tables with identical row and column labels. If your data have entire rows or columns missing because they were never reported by the raters, you must add a row or column of zeros to make the table square (see Example 6).

The results of this section are based on Fleiss, Levin, and Paik (2003). The kappa procedure also outputs the *Maximum Kappa* and *Maximum-Adjusted Kappa* statistics.

Kappa Estimation

Define the overall proportion of observed agreement, p_o, as

$$p_o = \sum_i p_{ii}$$

and the overall chance-expected proportion of agreement, p_e, as

$$p_e = \sum_i p_{i\cdot} p_{\cdot i}$$

Kappa is calculated as from p_o and p_e as

$$\hat{\kappa} = \frac{p_o - p_e}{1 - p_e}$$

with asymptotic standard error

$$\widehat{SE}_{\hat{\kappa}} = \frac{\sqrt{A + B - C}}{(1 - p_e)\sqrt{N}}$$

where

$$A = \sum_i p_{ii}[1 - (p_{i\cdot} + p_{\cdot i})(1 - \hat{\kappa})]^2$$

$$B = (1 - \hat{\kappa})^2 \sum_i \sum_{j \neq i} p_{ij}\left(p_{\cdot i} + p_{j\cdot}\right)^2$$

$$C = [\hat{\kappa} - p_e(1 - \hat{\kappa})]^2$$

An approximate $100(1 - \alpha)\%$ confidence interval for κ is

$$\hat{\kappa} - z_{\alpha/2}\widehat{SE}_{\hat{\kappa}} \leq \kappa \leq \hat{\kappa} + z_{\alpha/2}\widehat{SE}_{\hat{\kappa}}$$

Kappa Hypothesis Test

To test the null hypothesis that $\kappa = 0$, the standard error of kappa under the null hypothesis is calculated as

$$\widehat{SE}_{\hat{\kappa}_0} = \frac{1}{(1-p_e)\sqrt{N}}\sqrt{p_e + p_e^2 - \sum_i p_{i\cdot}p_{\cdot i}\,(p_{i\cdot} + p_{\cdot i})}$$

and the kappa test statistic, z_κ, with asymptotic standard normal distribution is

$$z_\kappa = \frac{\hat{\kappa}}{\widehat{SE}_{\hat{\kappa}_0}}$$

Weighted Kappa Estimation

Weighted kappa should only be used when the rater categories are ordered (e.g. "Low", "Medium", "High" or 1, 2, 3, 4). The procedure applies weights to quantify relative distances between categories. These weights can be calculated as either linear or quadratic in NCSS.

The **linear weights** are calculated as

$$w_{ij} = 1 - \frac{|i-j|}{R-1}$$

with *R* = *C*. For a 4 × 4 table, the linear weight matrix would be

1.00	0.67	0.33	0.00
0.67	1.00	0.67	0.33
0.33	0.67	1.00	0.67
0.00	0.33	0.67	1.00

The **quadratic weights** are calculated as

$$w_{ij} = 1 - \frac{(i-j)^2}{(R-1)^2}$$

again, with *R* = *C*. For a 4 × 4 table, the quadratic weight matrix would be

1.00	0.89	0.56	0.00
0.89	1.00	0.89	0.56
0.56	0.89	1.00	0.89
0.00	0.56	0.89	1.00

Note that in both cases the weights for cells on the diagonal are equal to 1 and weights off the diagonal are between 0 and 1. Weighted kappa is the same as simple kappa when using a weight matrix with all diagonal weight elements equal to 1 and all off-diagonal weight elements equal to 0.

Using the cell weights, we can calculate the observed weighted proportion of agreement as

$$p_{o_w} = \sum_i \sum_j w_{ij} p_{ij}$$

and the overall chance-expected weighted proportion of agreement, p_e, as

$$p_{e_w} = \sum_i \sum_j w_{ij} p_{i\cdot} p_{\cdot j}$$

Further define

$$\bar{w}_{i\cdot} = \sum_j w_{ij} p_{\cdot j}$$

$$\bar{w}_{\cdot j} = \sum_i w_{ij} p_{i\cdot}$$

Weighted kappa is calculated as

$$\hat{\kappa}_w = \frac{p_{o_w} - p_{e_w}}{1 - p_{e_w}}$$

with asymptotic standard error

$$\widehat{SE}_{\hat{\kappa}_w} = \frac{\sqrt{A - B}}{(1 - p_{e_w})\sqrt{N}}$$

where

$$A = \sum_i \sum_j p_{ij}\left[w_{ij} - (\bar{w}_{i\cdot} + \bar{w}_{\cdot j})(1 - \hat{\kappa}_w)\right]^2$$

$$B = \left[\hat{\kappa}_w - p_{e_w}(1 - \hat{\kappa}_w)\right]^2$$

An approximate $100(1-\alpha)\%$ confidence interval for κ_w is

$$\hat{\kappa}_w - z_{\alpha/2}\widehat{SE}_{\hat{\kappa}_w} \le \kappa_w \le \hat{\kappa}_w + z_{\alpha/2}\widehat{SE}_{\hat{\kappa}_w}$$

Weighted Kappa Hypothesis Test

To test the null hypothesis that $\kappa_w = 0$, the standard error of weighted kappa under the null hypothesis is calculated as

$$\widehat{SE}_{\hat{\kappa}_{w_0}} = \frac{1}{(1 - p_{e_w})\sqrt{N}} \sqrt{\sum_i \sum_j p_{i\cdot} p_{\cdot j}\left[w_{ij} - (\bar{w}_{i\cdot} + \bar{w}_{\cdot j})\right]^2 - p_{e_w}^2}$$

and the weighted kappa test statistic, z_{κ_w}, with asymptotic standard normal distribution is

$$z_{\kappa_w} = \frac{\hat{\kappa}_w}{\widehat{SE}_{\hat{\kappa}_{w_0}}}$$

Maximum-Adjusted Kappa

If we define the overall chance-expected proportion of agreement, p_e, as before with

$$p_e = \sum_i p_{i\cdot} p_{\cdot i}$$

and

$$p_{max} = \sum_i \min(p_{i\cdot}, p_{\cdot i})$$

then the maximum kappa for a table with the observed marginal totals, $\hat{\kappa}_{max}$, can be calculated as

$$\hat{\kappa}_{max} = \frac{p_{max} - p_e}{1 - p_e}$$

The maximum-adjusted kappa statistic, $\hat{\kappa}_{max-adj}$, is calculated as

$$\hat{\kappa}_{max-adj} = \frac{\hat{\kappa}}{\hat{\kappa}_{max}}$$

Association and Correlation Statistics

Phi

A measure of association independent of the sample size. *Phi* ranges between 0 (no relationship) and 1 (perfect relationship). *Phi* was designed for 2 × 2 tables only. For larger tables, it has no upper limit and Cramer's V should be used instead. The formula is

$$\phi = \sqrt{\frac{\chi_P^2}{N}}$$

Cramer's V

A measure of association independent of sample size. This statistic is a modification of the Phi statistic so that it is appropriate for larger than 2 × 2 tables. V ranges between 0 (no relationship) and 1 (perfect relationship).

$$V = \sqrt{\frac{\phi^2}{\min(R, C)}}$$

Pearson's Contingency Coefficient

A measure of association independent of sample size. It ranges between 0 (no relationship) and 1 (perfect relationship). For any particular table, the maximum possible depends on the size of the table (a 2 × 2 table has a maximum of 0.707), so it should only be used to compare tables with the same dimensions. The formula is

$$C = \sqrt{\frac{\chi_P^2}{\chi_P^2 + N}}$$

Tschuprow's T

A measure of association independent of sample size. This statistic is a modification of the Phi statistic so that it is appropriate for larger than 2 × 2 tables. T ranges between 0 (no relationship) and 1 (perfect relationship), but 1 is only attainable for square tables. The formula is

$$T = \sqrt{\frac{\phi^2}{\sqrt{(\mathrm{R}-1)(C-1)}}}$$

Lambda A - Rows dependent

This is a measure of association for cross tabulations of nominal-level variables. It measures the percentage improvement in predictability of the dependent variable (row variable or column variable), given the value of the other variable (column variable or row variable). The formula is

$$\lambda_a = \frac{\sum_i \max\left(O_{ij}\right) - \max(n_{i\cdot})}{N - \max(n_{i\cdot})}$$

Lambda B - Columns dependent

See Lambda A above. The formula is

$$\lambda_a = \frac{\sum_j \max\left(O_{ij}\right) - \max\left(n_{\cdot j}\right)}{N - \max\left(n_{\cdot j}\right)}$$

Symmetric Lambda

This is a weighted average of the *Lambda A* and *Lambda B* above. The formula is

$$\lambda = \frac{\sum_i \max\left(O_{ij}\right) + \sum_j \max\left(O_{ij}\right) - \max(n_{i\cdot}) - \max\left(n_{\cdot j}\right)}{2N - \max(n_{i\cdot}) - \max\left(n_{\cdot j}\right)}$$

Kendall's tau-B

This is a measure of correlation between two ordinal-level (rankable) variables. It is most appropriate for square tables. To compute this statistic, you first compute two values, *P* and *Q*, which represent the number of concordant and discordant pairs, respectively. The formula is

$$\tau_b = \frac{P - \mathrm{Q}}{N(N-1)/2}$$

Kendall's tau-B (with correction for ties)

This is the same as the above, except a correction is made for the case when ties are found in the data.

Kendall's tau-C

This is used in the case where the number of rows does not match the number of columns. The formula is

$$\tau_c = \frac{P - \mathrm{Q}}{N^2(\min(R,C) - 1)/2\min(R,C)}$$

Gamma

This is another measure based on concordant (P) and discordant (Q) pairs. The formula is

$$\gamma = \frac{P - \mathrm{Q}}{P + \mathrm{Q}}$$

Multiple Comparisons (2 × *k* Tables)

When one variable is nominal, the other is binary, and the overall test for independence is significant, you may wish to determine which of the group proportions are different from the others using a multiple comparison procedure. Two general types of multiple comparisons are often used: pairwise multiple comparisons and comparison of each treatment group with a control group. Both of these multiple comparison procedures are available in this procedure for binary data.

Another approach that has been promoted over the years is to ignore the non-normal nature of the data and use multiple comparison procedures that are available for comparing group means of normal data. The binary data values are recoded to numeric values of 0's and 1's. The group means of such data are the group proportions. For large samples, the central-limit theorem may be used to adopt either the Tukey-Kramer or the Dunnett multiple comparison procedure. For example, Dunnett's procedure is evaluated in Chuang-Stein and Tong (1995) who conclude that it performs reasonably well even for moderate sample sizes. Since these procedures are available in NCSS, we let you use those procedures if you desire. This section documents procedures that are specifically designed for binary data.

Pairwise Multiple Comparisons of the Difference of Binomial Proportions

This section uses the results presented in Agresti, Bini, Bertaccini, and Ryu (2008). They recommend constructing simultaneous confidence intervals using the Studentized range distribution with a score statistic. They also report that the adjusted Wald test performs reasonably well, except when the proportions are close to 0. We refer you to this article and to the references they give for further details. We will summarize the problem here.

Suppose we have binary responses in each of G independent groups. Let y_g represent a binomial variate based on n_g observations from group g (g = 1, …, G). Let p_g represent the population proportion of successes in group g. The maximum likelihood estimate of p_g is

$$\hat{p}_g = \frac{y_g}{n_g}$$

The Wald test of the difference between two groups, a and b, *is* constructed using the ratio of a difference divided by its standard error. In this case, the test statistic is

$$z_{a-b} = \frac{\hat{p}_a - \hat{p}_b}{SE_{a-b}}$$

where

$$SE_{a-b} = \sqrt{\frac{\hat{p}_a(1-\hat{p}_a)}{n_a} + \frac{\hat{p}_b(1-\hat{p}_b)}{n_b}}$$

Adjusted Pairwise Multiple Comparisons of Proportion Differences

This section discusses the *adjusted Tukey-type multiple comparisons of proportion differences*. We will present both a p-value and a set of simultaneous confidence intervals for all pairwise differences.

The p-value depends on the distribution of z_{a-b}. For the Tukey-type multiple comparison test, Agresti *et al.* (2008) assume that z_{a-b} follows the Studentized range distribution with infinite error degrees of freedom. We represent the critical value as $q_{G,\infty,\alpha}$.

The Tukey-type simultaneous confidence interval for the difference in proportions is

$$\hat{p}_a - \hat{p}_b \pm \frac{q_{G,\infty,1-\alpha}}{\sqrt{2}} SE_{a-b}$$

Many authors have shown that for various reasons, adding a small amount ε to each cell of the 2-by-G table of counts increases the accuracy of the confidence interval. Agresti and Caffo (2000) showed that setting ε to one works well. We call this the *adjusted Tukey-type* simultaneous confidence intervals. These $G(G-1)/2$ intervals have a family wise error rate (FWER) of $1-\alpha$.

Score-Type Pairwise Multiple Comparisons of Proportion Differences

This section discusses the *score-type multiple comparisons of proportion differences.* We will present both a p-value and a set of simultaneous confidence interval for all pairwise differences. The details of this method are given in Agresti *et al.* (2008).

The p-value matches the p-value obtained from the adjusted Tukey-type test described above.

The score-type simultaneous confidence interval for the difference in proportions is computed by inverting the score test having the test statistic

$$w_s = \left(\frac{(\hat{p}_a - \hat{p}_b) - \delta_{a,b,0}}{SE_{\widetilde{a-b}}} \right)^2$$

where

$$SE_{\widetilde{a-b}} = \sqrt{\frac{\tilde{p}_a(1-\tilde{p}_a)}{n_a} + \frac{\tilde{p}_b(1-\tilde{p}_b)}{n_b}}$$

Here, $\tilde{p}_a$ and $\tilde{p}_b$ are the maximum likelihood estimates of p_a and p_b constrained so that $\tilde{p}_a - \tilde{p}_b = \delta_{a,b,0}$.

A critical value, C, is computed from the Studentized range distribution using

$$C = \left(\frac{q_{G,\infty,1-\alpha}}{\sqrt{2}} \right)^2$$

The test inversion is accomplished as follows. A fine grid of about 2000 possible values for $\delta_{a,b,0}$ is obtained and w_s is calculated for each of these values. The minimum value of $\delta_{a,b,0}$ for which $w_s < C$ becomes the lower confidence limit and the maximum value of $\delta_{a,b,0}$ for which $w_s < C$ become the upper confidence limit.

Pairwise Multiple Comparisons of Proportion Differences using the Angular Transformation

This section discusses the *multiple comparisons of proportion differences using the angular transformation.* The details of this method are given in Zar (2010), page 557.

The p-value is found using modified versus of the Tukey-type test described above. The modification is to replace the two proportions that are differenced with the Freeman and Tukey (1950) transformed values. This transformation is as follows

$$p' = \frac{1}{2}\left[\arcsin\sqrt{\frac{y}{n+1}} + \arcsin\sqrt{\frac{y+1}{n+1}} \right]$$

The *arcsin* result is in degrees (rather than radians). The standard error of the difference of the transformed proportions is

$$SE_A = \sqrt{\frac{410.35}{n_a} + \frac{410.35}{n_b}}$$

The p-value of the Wald test is obtained from the Studentized range distribution with infinite error degrees of freedom. The Wald test is given by

$$W_{a-b} = \frac{p'_a - p'_b}{SE_A}$$

No confidence intervals are available in the original proportion scale.

Score-Type Pairwise Multiple Comparisons of Odds Ratios

This section discusses the *score-type multiple comparisons of odds ratios*. We will present both a p-value and a set of simultaneous confidence interval for all pairwise odds ratios. The details of this method are given in Agresti *et al.* (2008).

Assume that the group subscripts are *i* and *j*. The score-type simultaneous confidence interval for the odds ratio of the two group proportions is computed by inverting the score test having the test statistic

$$z_{ij,0}^2(\psi_{ij,0}) = \frac{\{n_i(\hat{p}_i - \tilde{p}_i)\}^2}{n_i\tilde{p}_i(1-\tilde{p}_i)} + \frac{\{n_j(\hat{p}_j - \tilde{p}_j)\}^2}{n_j\tilde{p}_j(1-\tilde{p}_j)}$$

Here, $\tilde{p}_i$ and $\tilde{p}_j$ are the maximum likelihood estimates of p_i and p_j constrained so that $\psi_{ij,0} = OR(\tilde{p}_i, \tilde{p}_j)$

where

$$\mathrm{OR}(\tilde{p}_i, \tilde{p}_j) = \frac{\mathrm{Odds}(\tilde{p}_i)}{\mathrm{Odds}(\tilde{p}_j)} = \frac{\dfrac{\tilde{p}_i}{1-\tilde{p}_i}}{\dfrac{\tilde{p}_j}{1-\tilde{p}_j}}$$

A critical value, *C*, is computed from the Studentized range distribution using $C = \left(\frac{q_{G,\infty,1-\alpha}}{\sqrt{2}}\right)^2$.

The test inversion is accomplished as follows. A fine grid of about 2000 possible values for $\psi_{ic,0}$ is obtained and $z_{ij,0}^2(\psi_{ij,0})$ is calculated for each of these values. The minimum value of $\psi_{ij,0}$ for which $z_{ij,0}^2(\psi_{ij,0}) < C$ becomes the lower confidence limit and the maximum value of $\psi_{ij,0}$ for which $z_{ij,0}^2(\psi_{ij,0}) < C$ become the upper confidence limit.

Many authors have shown that for various reasons, adding a small amount ε to each cell of the 2-by-G table of counts increases the accuracy of the confidence interval. Agresti and Caffo (2000) indicated that setting ε to one works well. NCSS gives you the option to set ε from 0 to 5.

Multiple Comparisons of Treatment Proportions with a Control

This section uses results similar to those presented in Chuang-Stein and Tong (1995) and Zar (2010). We construct simultaneous confidence intervals using Dunnett's distribution with infinite degrees of freedom and using an adjusted Wald statistic.

Suppose we have binary responses in each of *G* independent groups, with the control group designated as group *c*. Let y_g represent a binomial variate based on n_g observations from group *g* (*g* = 1, …, *G*). Let p_g represent the population proportion of successes in group *g*. The maximum likelihood estimate of p_g is

$$\hat{p}_g = \frac{y_g}{n_g}$$

The Wald test of the difference between a treatment group and the control group, *a* and *c*, *is* constructed using the ratio of a difference divided by its standard error. In this case, the test statistic is

$$z_{ac} = \frac{\hat{p}_a - \hat{p}_c}{SE_{ac}}$$

where

$$SE_{ac} = \sqrt{\frac{\hat{p}_a(1-\hat{p}_a)}{n_a} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_c}}$$

Adjusted Dunnett-Type Multiple Comparisons of Proportion Differences

This section discusses the *adjusted Dunnett-type multiple comparisons of proportion differences.* We will present both a p-value and a set of simultaneous confidence intervals for the $G - 1$ differences.

The p-value depends on the distribution of z_{ac}. For the Dunnett-type multiple comparison test, we use the Dunnett's distribution with infinite degrees. We represent this critical value as $d_{G,\infty,\alpha}$. Note that this is an approximation to the actual distribution which depends on the correlation among groups. As some authors have pointed out, there is little difference between the exact calculation based on the multivariate normal distribution and specified correlation matrix (which is assumed to be known), and this procedure.

The Dunnett-type simultaneous confidence interval for the difference in proportions is

$$\hat{p}_a - \hat{p}_c \pm d_{G,\infty,1-\alpha} SE_{ac}$$

Agresti and Caffo (2000) showed that for various reasons, adding a small amount ε to each cell of the 2-by-G table of counts increases the accuracy of the confidence interval. They showed that setting ε to one works well. We call this the *adjusted Dunnett-type* simultaneous confidence intervals. These $G - 1$ simultaneous intervals have a family wise error rate (FWER) of $1 - \alpha$.

These intervals can be one-sided or two-sided.

Score-Type Many-to-One Multiple Comparisons of Odds Ratios

This section discusses the *score-type multiple comparisons of odds ratios.* We will present both a p-value and a set of simultaneous confidence interval for all pairwise odds ratios. The details of this method are given in Agresti *et al.* (2008).

Assume that the control group has subscript *c* and each of the *G* – 1 treatment groups has subscript *i*. The score-type simultaneous confidence interval for the difference in proportions is computed by inverting the score test having the test statistic

$$z_{ic,0}^2(\psi_{ic,0}) = \frac{\{n_i(\hat{p}_i - \tilde{p}_i)\}^2}{n_i\tilde{p}_i(1-\tilde{p}_i)} + \frac{\{n_c(\hat{p}_c - \tilde{p}_c)\}^2}{n_c\tilde{p}_c(1-\tilde{p}_c)}$$

Here, $\tilde{p}_i$ and $\tilde{p}_c$ are the maximum likelihood estimates of p_i and p_c constrained so that $\psi_{ic,0} = OR(\tilde{p}_i, \tilde{p}_c)$

where

$$\mathrm{OR}(\tilde{p}_i, \tilde{p}_c) = \frac{\mathrm{Odds}(\tilde{p}_i)}{\mathrm{Odds}(\tilde{p}_c)} = \frac{\frac{\tilde{p}_i}{1-\tilde{p}_i}}{\frac{\tilde{p}_c}{1-\tilde{p}_c}}$$

A critical value, *C*, is computed from Dunnett's distribution, $d_{G,\infty,\alpha}$, using $C = \left(d_{G,\infty,1-\alpha}\right)^2$.

The test inversion is accomplished as follows. A fine grid of about 2000 possible values for $\psi_{ic,0}$ is obtained and $z_{ic,0}^2(\psi_{ic,0})$ is calculated for each of these values. The minimum value of $\psi_{ic,0}$ for which $z_{ic,0}^2(\psi_{ic,0}) < C$ becomes the lower confidence limit and the maximum value of $\psi_{ic,0}$ for which $z_{ic,0}^2(\psi_{ic,0}) < C$ become the upper confidence limit.

Many authors have shown that for various reasons, adding a small amount ε to each cell of the 2-by-G table of counts increases the accuracy of the confidence interval. Agresti and Caffo (2000) indicated that setting ε to one works well. NCSS gives you the option to set ε from 0 to 5.

Data Structure

You may use either summarized or non-summarized data for this procedure. Typically, you will use data columns of categorical data. If you want to perform crosstab analysis on numeric data, the data must be grouped into categories before a table can be created. This is best accomplished by using an *If-Then* or *Recode* transformation. You can also use this procedure's facility to categorize numeric data by checking "Create Other Row/Column Variables for Numeric Data" on the Variables tab.

The following are two example datasets that illustrate the type of data that can be analyzed using this procedure. The datasets are provided with the software. The first dataset "CrossTabs1" contains fictitious responses to a survey of 100 people in which respondents were asked about their weekly sugar intake and exercise. The data are in raw form. The second hypothetical dataset "McNemar" contains summarized responses from 23 individuals who were asked about their desire to purchase a certain home-improvement product before and after a sales demonstration.

CrossTabs1 dataset (subset)

Sugar	Exercise
High	Infrequent
Low	Frequent
High	Infrequent
High	Infrequent
High	Frequent
High	Infrequent
High	Infrequent
High	Infrequent
High	Infrequent
Low	Frequent

McNemar dataset

Before	After	Count
No	Yes	10
No	No	6
Yes	Yes	4
Yes	No	3

The data below are a subset of the *Real Estate Sales* database provided with the software. This (computer-simulated) data gives information including the selling price, the number of bedrooms, the total square footage (finished and unfinished), and the size of the lots for 150 residential properties sold during the last four months in two states. Only the first 6 of 150 observations are displayed here. The variables "Price", "TotalSqft", and "LotSize" would need to be categorized before they could be displayed in a table.

Resale dataset (subset)

State	Price	Bedrooms	TotalSqft	LotSize
Nev	260000	2	2042	10173
Nev	66900	3	1392	13069
Vir	127900	2	1792	7065
Nev	181900	3	2645	8484
Nev	262100	2	2613	8355

Missing Values

Missing values may be ignored or included in the table's counts, percentages, statistics, and tests. This is controlled on the procedure's Missing tab.

Two-Way Table Data Input

NCSS also allows you to input the contingency table data directly into the procedure without using the Data Window. To do this, select "Two-Way Table" for Type of Data Input.

Example 1 – 2 × 2 Contingency Table and Statistics from Raw Categorical Data

The data for this example are found in the "CrossTabs1" dataset. This dataset contains fictitious survey data from 100 individuals asked about their sugar intake and exercise. Notice that we have entered custom value orders for the columns in the dataset so that the values will appear in the correct order. We use a 2 × 2 contingency table for this example so that all of the tests for row-column independence will be displayed.

Setup

To run this example, complete the following steps:

1 Open the CrossTabs1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **CrossTabs1** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**Sugar**
Column Variable(s)	**Exercise**
Reports Tab	
Show Individual Tables	**Checked**
Counts	**Checked**
Show Combined Table	**Checked**
Counts	**Checked**
Table Percentages	**Checked**
Row Percentages	**Checked**
Column Percentages	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Checked**
Pearson's Chi-Square Test	**Checked**
Yates' Continuity Corrected Chi-Square Test [2×2 Tables]	**Checked**
Likelihood Ratio Test	**Checked**
Fisher's Exact Test [2×2 Tables]	**Checked**
Association and Correlation Statistics	**Checked**
Plots Tab	
Show Bar Charts	**Checked**
Counts	**Checked**
Report Options (*in the Toolbar*)	
Variable Labels	**Column Names**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Sugar	Infrequent (Exercise)	Frequent (Exercise)	Total
Low	19	28	47
High	37	16	53
Total	56	44	100

Combined Table

Sugar		Infrequent (Exercise)	Frequent (Exercise)	Total
Low	Count	19	28	47
	% of Total	19.00%	28.00%	47.00%
	% within Row	40.43%	59.57%	100.00%
	% within Column	33.93%	63.64%	47.00%
High	Count	37	16	53
	% of Total	37.00%	16.00%	53.00%
	% within Row	69.81%	30.19%	100.00%
	% within Column	66.07%	36.36%	53.00%
Total	Count	56	44	100
	% of Total	56.00%	44.00%	100.00%
	% within Row	56.00%	44.00%	100.00%
	% within Column	100.00%	100.00%	100.00%

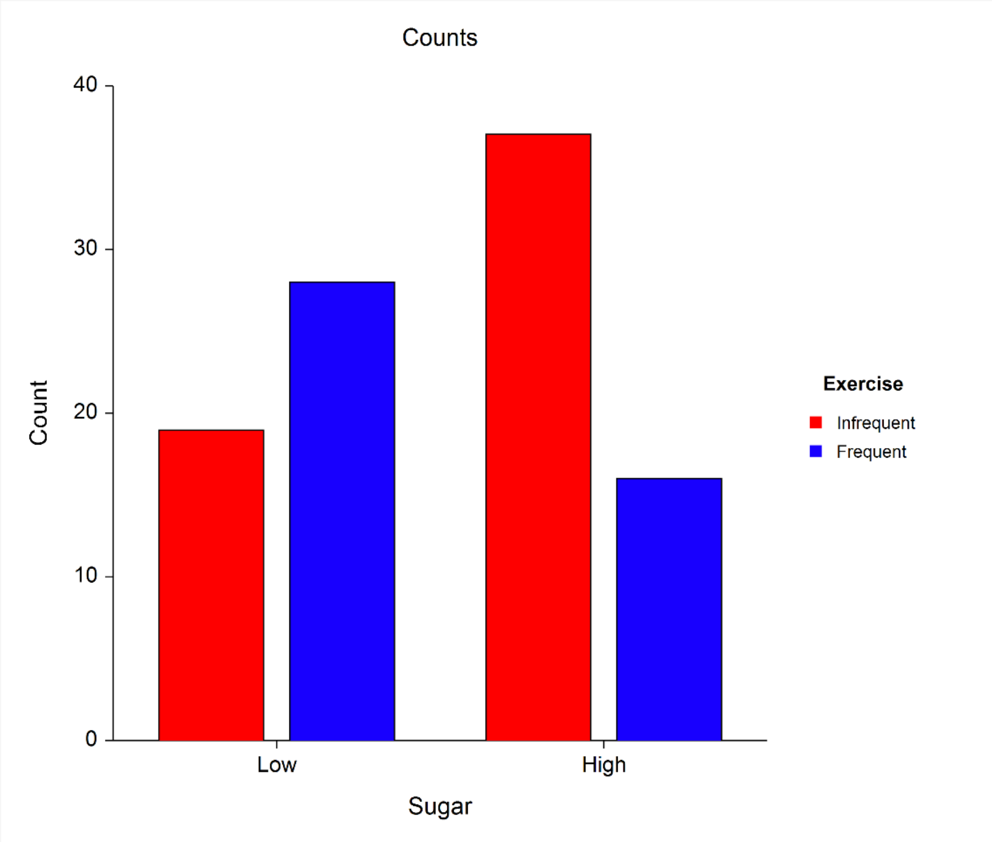
Tests for Row-Column Independence
(Sugar by Exercise)
H0: "Sugar" and "Exercise" are independent.
H1: "Sugar" and "Exercise" are associated (not independent).

Test	Type	Chi-Square Value	DF	Prob Level	Reject H0 at α = 0.05?
Pearson's Chi-Square	2-Sided	8.7299	1	0.00313	Yes
Yates' Cont. Correction	2-Sided	7.5780	1	0.00591	Yes
Likelihood Ratio	2-Sided	8.8440	1	0.00294	Yes
Fisher's Exact	2-Sided			0.00458	Yes
Fisher's Exact (Lower)	1-Sided			0.00284	Yes
Fisher's Exact (Upper)	1-Sided			0.99926	No

Association and Correlation Statistics
(Sugar by Exercise)

Statistic	Value
Phi	0.2955
Cramer's V	0.2955
Pearson's Contingency Coefficient	0.2834
Tschuprow's T	0.2955
Lambda A .. Rows dependent	0.2553
Lambda B .. Columns dependent	0.2045
Symmetric Lambda	0.2308
Kendall's tau-B	-0.1479
Kendall's tau-B (with correction for ties)	-0.2955
Kendall's tau-C	-0.2928
Gamma	-0.5463

Plots Section
(Sugar by Exercise)



This report presents the individual contingency table of counts, a combined table with counts and percentages, the results of the various row-column independence tests, and various association and correlation statistics. A plot of the counts is also displayed. The Pearson's chi-square test results indicate that for these hypothetical data there is an association between a person's sugar intake and exercise frequency (p-value = 0.00313).

Example 2 – 3 × 4 Contingency Table and Statistics from Summarized Categorical Data

The data for this example are found in the "CrossTabs2" dataset. Notice that we have entered custom value orders for the column labeled Region so that the values will appear in the correct order.

Setup

To run this example, complete the following steps:

1 Open the CrossTabs2 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **CrossTabs2** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2a** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**Region**
Column Variable(s)	**Choice**
Frequency Variable	**Count**
Reports Tab	
Show Individual Tables	**Unchecked**
Show Combined Table	**Checked**
Counts	**Checked**
Table Percentages	**Checked**
Row Percentages	**Checked**
Column Percentages	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Checked**
Pearson's Chi-Square Test	**Checked**
Yates' Continuity Corrected Chi-Square Test [2×2 Tables]	**Checked**
Likelihood Ratio Test	**Checked**
Fisher's Exact Test [2×2 Tables]	**Checked**
Plots Tab	
Show Bar Charts	**Checked**
Counts	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Combined Table

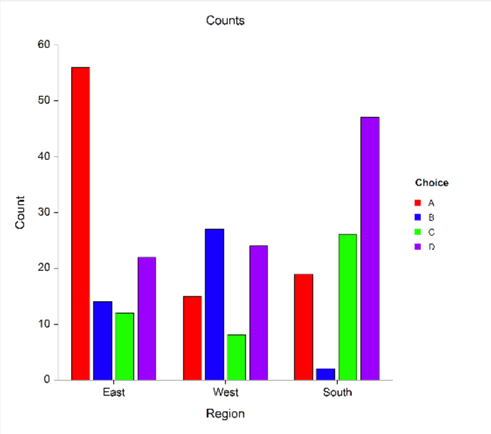
Region		Choice A	B	C	D	Total
East	Count	56	14	12	22	104
	% of Total	20.59%	5.15%	4.41%	8.09%	38.24%
	% within Row	53.85%	13.46%	11.54%	21.15%	100.00%
	% within Column	62.22%	32.56%	26.09%	23.66%	38.24%
West	Count	15	27	8	24	74
	% of Total	5.51%	9.93%	2.94%	8.82%	27.21%
	% within Row	20.27%	36.49%	10.81%	32.43%	100.00%
	% within Column	16.67%	62.79%	17.39%	25.81%	27.21%
South	Count	19	2	26	47	94
	% of Total	6.99%	0.74%	9.56%	17.28%	34.56%
	% within Row	20.21%	2.13%	27.66%	50.00%	100.00%
	% within Column	21.11%	4.65%	56.52%	50.54%	34.56%
Total	Count	90	43	46	93	272
	% of Total	33.09%	15.81%	16.91%	34.19%	100.00%
	% within Row	33.09%	15.81%	16.91%	34.19%	100.00%
	% within Column	100.00%	100.00%	100.00%	100.00%	100.00%

Tests for Row-Column Independence
(Region by Choice)
H0: "Region" and "Choice" are independent.
H1: "Region" and "Choice" are associated (not independent).

Test	Type	Chi-Square Value	DF	Prob Level	Reject H0 at α = 0.05?
Pearson's Chi-Square†	2-Sided	75.3662	6	0.00000	Yes
Yates' Cont. Correction*					
Likelihood Ratio	2-Sided	75.0616	6	0.00000	Yes
Fisher's Exact*					

† WARNING: At least one cell had a value less than 5.
* Test computed only for 2×2 tables.

Plots Section
(Region by Choice)



This report presents the results from the summarized data. The Pearson's chi-square test results indicate that the row and column variables are not independent (p-value = 0.00000), but there is a sample size warning that should be considered. Note that Fisher's Exact Test and Yates' Continuity Correction are not reported because this is not a 2 × 2 table.

An alternate way to enter this summarized data is to set **Type of Data Input** to **Two-Way Table** and enter the titles, labels, and counts into the two-way table on the input window.

4 Modify the Data Input Type

- The settings for this section are listed below and are stored in the **Example 2b** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Two-Way Table**
Rows	**3**
Columns	**4**

5 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

The output will be exactly the same as that displayed above.

Example 3 – 6 × 4 Contingency Table and Statistics from Raw Numeric Data

The real estate data for this example are found in the “Resale” dataset. We will use the crosstabs procedure to create a table with city as the row variable and price groups as the column variable. The software will summarize the continuous price variable for us using a list of price group boundaries. We will use the column labels and value labels in the dataset to make the data easier to interpret in the reports.

Setup

To run this example, complete the following steps:

1 Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 3** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**City**
Column Variable(s)	**<Empty>**
Create Other Column Variables from Numeric Data	**Checked**
Numeric Variable(s) to Categorize for Use in Table Rows	**Price**
Group Numeric Data into Categories Using	**List of Interval Upper Limits**
List	**100000 200000 300000**
Reports Tab	
Show Individual Tables	**Checked**
Counts	**Checked**
Row Percentages	**Checked**
Show Combined Table	**Unchecked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Checked**
Pearson's Chi-Square Test	**Checked**
Yates' Continuity Corrected Chi-Square Test [2×2 Tables]	**Checked**
Likelihood Ratio Test	**Checked**
Fisher's Exact Test [2×2 Tables]	**Checked**

Report Options (*in the Toolbar*)
Variable Labels .... **Column Labels**
Data Labels.. **Value Labels**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Community	Up To 100000	100000 To 200000	200000 To 300000	Over 300000	Total
	Sales Price				
Silverville	5	12	4	6	27
Los Wages	18	16	9	6	49
Red Gulch	5	3	3	1	12
Politicville	5	11	10	1	27
Senate City	6	12	4	2	24
Congresstown	2	4	2	3	11
Total	41	58	32	19	150

Row Percentages Table

Community	Up To 100000	100000 To 200000	200000 To 300000	Over 300000	Total
	Sales Price				
Silverville	18.52%	44.44%	14.81%	22.22%	100.00%
Los Wages	36.73%	32.65%	18.37%	12.24%	100.00%
Red Gulch	41.67%	25.00%	25.00%	8.33%	100.00%
Politicville	18.52%	40.74%	37.04%	3.70%	100.00%
Senate City	25.00%	50.00%	16.67%	8.33%	100.00%
Congresstown	18.18%	36.36%	18.18%	27.27%	100.00%
Total	27.33%	38.67%	21.33%	12.67%	100.00%

Tests for Row-Column Independence
(Community by Sales Price)
H0: "Community" and "Sales Price" are independent.
H1: "Community" and "Sales Price" are associated (not independent).

Test	Type	Chi-Square Value	DF	Prob Level	Reject H0 at α = 0.05?
Pearson's Chi-Square†	2-Sided	16.8045	15	0.33069	No
Yates' Cont. Correction*					
Likelihood Ratio	2-Sided	16.2412	15	0.36620	No
Fisher's Exact*					

† WARNING: At least one cell had an expected value less than 5.
* Test computed only for 2×2 tables.

This report presents the results from the data with the continuous price variable grouped into 4 categories. The Pearson's chi-square test results indicate that there is not enough evidence to conclude that the row and column variables are associated (p-value = 0.33069). There is an expected value warning that should be considered. Note that Fisher's Exact Test and Yates' Continuity Correction are not reported because this is not a 2 × 2 table.

Example 4 – Tests for Trend in Proportions (Validation using Armitage (1955))

The data for this example come from Table 1 of Armitage (1955) and are stored in the "Armitage" dataset. The dataset contains counts of tonsil sizes (+, ++, +++) from 1398 children aged 0-15 years along with and indicator or whether each child is a carrier or non-carrier of the bacteria Streptococcus pyogenes. On page 378, Armitage (1955) calculates the Cochran-Armitage chi-square test statistic for the alternative hypothesis of any trend to be 7.19 on 1 df with a p-value of 0.007. On page 383, Armitage (1955) calculates the Rank Correlation Test chi-square test statistic for the alternative hypothesis of any trend to be 6.83 on 1 df with a p-value of 0.009.

Setup

To run this example, complete the following steps:

1 Open the Armitage example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Armitage** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 4** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Row Variable(s)	**Strep**
Column Variable(s)	**Tonsils**
Frequency Variable	**Count**
Reports Tab	
Show Individual Tables	**Unchecked**
Show Combined Table	**Checked**
Counts	**Checked**
Column Percentages	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Unchecked**
Tests for Trend in Proportions (Cochran-Armitage) [2×k Tables]	**Checked**
Cochran-Armitage Test	**Checked**
Cochran-Armitage Test with Continuity Correction	**Checked**
Armitage Rank Correlation Test	**Checked**
Report Options (*in the Toolbar*)	
Variable Labels	**Column Labels**
Data Labels	**Value Labels**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Combined Table

Streptococcus pyogenes		Tonsil Size 1	2	3	Total
Non-carriers	Count	497	560	269	1326
	% within Column	96.32%	95.08%	91.81%	94.85%
Carriers	Count	19	29	24	72
	% within Column	3.68%	4.92%	8.19%	5.15%
Total	Count	516	589	293	1398
	% within Column	100.00%	100.00%	100.00%	100.00%

Cochran-Armitage Trend Test
(Streptococcus pyogenes by Tonsil Size)
H0: p(1) = p(2) = p(3) = ... = p(k)

Alternative Hypothesis*	Numerator	Standard Error	Z	Prob Level	Reject H0 at α = 0.05?
H1: Increasing Trend	16.48498	6.1467	2.6819	0.00366	Yes
H1: Decreasing Trend	16.48498	6.1467	2.6819	0.99634	No
H1: Any Trend	16.48498	6.1467	2.6819	0.00732	Yes

* Trend is based on % within Column for Streptococcus pyogenes = "Carriers".

Cochran-Armitage Trend Test with Continuity Correction
(Streptococcus pyogenes by Tonsil Size)
H0: p(1) = p(2) = p(3) = ... = p(k)

Alternative Hypothesis*	Numerator†	Standard Error	Z	Prob Level	Reject H0 at α = 0.05?
H1: Increasing Trend	15.98498	6.1467	2.6006	0.00465	Yes
H1: Decreasing Trend	15.98498	6.1467	2.6006	0.99535	No
H1: Any Trend	15.98498	6.1467	2.6006	0.00931	Yes

* Trend is based on % within Column for Streptococcus pyogenes = "Carriers".
† Continuity Correction Factor (Δ/2) = 0.5

Armitage Rank Correlation Trend Test
(Streptococcus pyogenes by Tonsil Size)
H0: p(1) = p(2) = p(3) = ... = p(k)

Alternative Hypothesis*	Numerator S	Standard Error	Z	Prob Level	Reject H0 at α = 0.05?
H1: Increasing Trend	16229	6208.3460	2.6141	0.00447	Yes
H1: Decreasing Trend	16229	6208.3460	2.6141	0.99553	No
H1: Any Trend	16229	6208.3460	2.6141	0.00895	Yes

* Trend is based on % within Column for Streptococcus pyogenes = "Carriers".

The reported alternative hypotheses correspond to the trend in proportions for the second row (Strep = “Carriers”). The two-sided Cochran-Armitage test confirms that the carrier rate (% within Column for Strep = “Carriers”) does, in fact, change with the tonsil size (Z = 2.6819 and p-value = 0.00732). The continuity corrected test (p-value = 0.00931) and Armitage rank correlation test (Z = 2.6141 and p-value = 0.00895) show similar results. These test results match exactly those given in Armitage (1955) if we note that Z^2 = Chi-Square on 1 df such that 2.6819^2=7.1926 and 2.6141^2=6.8335.

Example 5 – McNemar Test

The data for this example are found in the "McNemar" dataset. This hypothetical data contains summarized responses from 23 individuals who were asked about their desire to purchase a certain home-improvement product before and after a sales demonstration.

Setup

To run this example, complete the following steps:

1 Open the McNemar example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **McNemar** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 5** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**Before**
Column Variable(s)	**After**
Frequency Variable	**Count**
Reports Tab	
Show Individual Tables	**Checked**
Counts	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Unchecked**
McNemar Test [k×k Tables]	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Before	After: No	Yes	Total
No	6	10	16
Yes	3	4	7
Total	9	14	23

McNemar Test
(Before by After)
H0: P12 = P21
H1: P12 ≠ P21

Test	Type	Chi-Square Value	DF	Prob Level	Reject H0 at α = 0.05?
Asymptotic Chi-Square	2-Sided	3.7692	1	0.05220	No
Binomial Exact	2-Sided			0.09229	No

Both the Asymptotic Chi-Square (p-value = 0.05220) and Binomial Exact (p-value = 0.09229) McNemar Tests indicate that there is not enough evidence to reject the null hypothesis.

Example 6 – Kappa Test for Inter-Rater Agreement from Summarized Data (Validation using Fleiss, Levin, and Paik (2003))

Fleiss, Levin, and Paik (2003) present a hypothetical example on pages 598-608 in which 100 subjects are diagnosed independently by two raters and placed into 1 of 3 categories: Psychotic, Neurotic, and Organic. The summarized data are contained in the "KappaFleiss" dataset. In this example dataset, please note that a custom value order has been entered so that the table categories appear in the same order as in the book.

They compute a kappa value of 0.68, a null standard error of 0.076, and a *z*-value of 8.95 for testing the null hypothesis that kappa = 0. They also compute the asymptotic standard error for computing confidence intervals to be 0.087.

Setup

To run this example, complete the following steps:

1 Open the KappaFleiss example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **KappaFleiss** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 6a** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**Rater_A**
Column Variable(s)	**Rater_B**
Frequency Variable	**Count**
Reports Tab	
Show Individual Tables	**Checked**
Counts	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Unchecked**
Kappa and Weighted Kappa Tests for Inter-Rater Agreement [k×k Tables]	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Rater_A \ Rater_B	Psychotic	Neurotic	Organic	Total
Psychotic	75	1	4	80
Neurotic	5	4	1	10
Organic	0	0	10	10
Total	80	5	15	100

Kappa Estimation
(Rater_A by Rater_B)

Statistic	Value	Asymptotic Std. Error	95% Lower Conf. Limit	95% Upper Conf. Limit
Kappa	0.6765	0.0877	0.5046	0.8484
Weighted Kappa (Linear)	0.7222	0.0843	0.5570	0.8874
Weighted Kappa (Quadratic)	0.7553	0.0867	0.5854	0.9253
Maximum-Adjusted Kappa*	0.7931			

* Maximum Kappa with the Observed Marginal Totals = 0.8529

Kappa Hypothesis Tests
(Rater_A by Rater_B)
H0: Kappa = 0
H1: Kappa > 0 (One-Sided)
H1: Kappa ≠ 0 (Two-Sided)

Test	Value	Asymptotic Std. Error under H0	Z	One-Sided Prob Level	Two-Sided Prob Level
Kappa	0.6765	0.0762	8.8791	0.00000	0.00000
Weighted Kappa (Linear)	0.7222	0.0879	8.2201	0.00000	0.00000
Weighted Kappa (Quadratic)	0.7553	0.0989	7.6335	0.00000	0.00000

The results from **NCSS** match Fleiss, Levin, and Paik (2003) with some small differences due to rounding (most pronounced in the *z*-value). The authors used rounded values in hand calculations to arrive at their results. **NCSS** uses full precision in all calculations.

The weighted kappa results are completely disregarded here because the rating categories have no inherent order. Weighted kappa should be ignored in this case.

An alternate way to enter this summarized data is to set **Type of Data Input** to **Two-Way Table** and enter the titles, labels, and counts into the two-way table on the input window.

4 Modify the Data Input Type

- The settings for this section are listed below and are stored in the **Example 6b** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Two-Way Table**
Rows	**3**
Columns	**3**

5 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

The output will be exactly the same as that displayed above.

Example 7 – Weighted Kappa Test from Raw Data with Missing Cell Combinations

The data for this example are contained in the "WeightedKappa" dataset. This dataset contains independent ratings by 2 raters on 12 individuals. Each individual was scored on a scale of 1 to 5. The goal is to determine how closely the raters' scores agree. Weighted kappa is appropriate in this case because the categories are ordinal, meaning that the categories have magnitude with natural ordering.

The problem with the data in this dataset, however, is that there are no cases where Rater 1 scored an individual as "3" and there are no cases where Rater 2 scored and individual as "2". This results in a contingency table that is square but does not have identical row and column categories. The second part of the example will show you how to modify the data by adding zeros appropriately so that kappa can be computed without having to go through the added step of summarizing the data first.

Setup

To run this example, complete the following steps:

1 Open the WeightedKappa example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **WeightedKappa** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 7a** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**Rater_1**
Column Variable(s)	**Rater_2**
Reports Tab	
Show Individual Tables	**Checked**
Counts	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Unchecked**
Kappa and Weighted Kappa Tests for Inter-Rater Agreement [k×k Tables]	**Checked**
Report Options (*in the Toolbar*)	
Variable Labels	**Column Labels**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Rater 1	Rater 2: 1	3	4	5	Total
1	3	0	0	0	3
2	0	3	0	0	3
4	0	1	2	0	3
5	0	0	1	2	3
Total	3	4	3	2	12

The number of rows with at least one missing value is 1

Kappa Estimation
(Rater 1 by Rater 2)

Not Calculated: Kappa statistics are only calculated for square k×k tables with identical row and column categories.

Kappa Hypothesis Tests
(Rater 1 by Rater 2)

Not Calculated: Kappa tests are only calculated for square k×k tables with identical row and column categories.

This report indicates that it is not possible to calculate the kappa test statistic for this table since the table categories are not identical for rows and columns. There is no row category "3" for Rater 1 and no column category "2" for Rater 2. This is a result of the fact that some categories were not applied by each rater. We could re-enter this summarized data into the data table, including missing rows and columns with assigned counts of "0", but there is an easier way. Simply add a new variable "Count" to the dataset and give each row a count value of "1". In the row immediately after the last data row, enter the values for each rater that were never observed (i.e. in the first empty row at the end of the dataset, enter "3" under Rater_1 and "2" under Rater_2) and assign those a count value of "0". In the example dataset we created new variables **RaterMod_1**, **RaterMod_2**, and **Count** to illustrate this principle, but in your dataset you do not necessarily have to create new rater variables; you can just add the rows with count = 0 starting at the first empty row. We will now run the analysis using the modified variables to get the kappa results.

4 Modify the variables

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 7b** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**RaterMod_1**
Column Variable(s)	**RaterMod_2**
Frequency Variable	**Count**

5 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Rater 2
Rater 1

	1	2	3	4	5	Total
1	3	0	0	0	0	3
2	0	0	3	0	0	3
3	0	0	0	0	0	0
4	0	0	1	2	0	3
5	0	0	0	1	2	3
Total	3	0	4	3	2	12

Kappa Estimation
(Rater 1 by Rater 2)

Statistic	Value	Asymptotic Std. Error	95% Lower Conf. Limit	95% Upper Conf. Limit
Kappa	0.5000	0.1429	0.2199	0.7801
Weighted Kappa (Linear)	0.7561	0.0968	0.5664	0.9458
Weighted Kappa (Quadratic)	0.9057	0.0481	0.8114	0.9999
Maximum-Adjusted Kappa*	0.8333			

* Maximum Kappa with the Observed Marginal Totals = 0.6000

Kappa Hypothesis Tests
(Rater 1 by Rater 2)
H0: Kappa = 0
H1: Kappa > 0 (One-Sided)
H1: Kappa ≠ 0 (Two-Sided)

Test	Value	Asymptotic Std. Error under H0	Z	One-Sided Prob Level	Two-Sided Prob Level
Kappa	0.5000	0.1173	4.2640	0.00001	0.00002
Weighted Kappa (Linear)	0.7561	0.1952	3.8725	0.00005	0.00011
Weighted Kappa (Quadratic)	0.9057	0.2856	3.1708	0.00076	0.00152

The counts table now has identical row and column categories with the exact same counts as the previous table. A row and column of zeros have been added to make the table suitable for the calculation of the kappa and weighted kappa statistics. It is appropriate to consider the weighted kappa statistic in this case because the data are comprised of ordered scores. A weighted kappa value of 0.7561 indicates moderate-to-high agreement between the raters. If we did not take into account the ordinal nature of the data and looked at the simple kappa statistic, we would conclude a much lower association of 0.5000. This demonstrates the importance of using weighted kappa when it is appropriate.

Example 8 – Data Summary Report

The data summary report was designed for situations in which you want to transfer a summarized table to another program. This format creates a vertical listing of the counts in a format that is easy to copy and paste into another **NCSS** dataset or into other programs. The data for this example are found in the "Resale" dataset.

Setup

To run this example, complete the following steps:

1 Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 8** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Columns in the Database**
Row Variable(s)	**City**
Column Variable(s)	**<Empty>**
Create Other Column Variables from Numeric Data	**Checked**
Numeric Variable(s) to Categorize for Use in Table Rows	**Price**
Group Numeric Data into Categories Using	**List of Interval Upper Limits**
List	**100000 200000 300000**
Reports Tab	
Data Summary Report	**Checked**
Show Individual Tables	**Checked**
Counts	**Checked**
Tests for Row-Column Independence (Pearson's Chi-Square, Fisher's Exact)	**Unchecked**
Report Options (*in the Toolbar*)	
Variable Labels	**Column Names**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Data Summary Report

Price	City	Count
Up To 100000	1	5
Up To 100000	2	18
Up To 100000	3	5
Up To 100000	4	5
Up To 100000	5	6
Up To 100000	6	2
100000 To 200000	1	12
100000 To 200000	2	16
100000 To 200000	3	3
100000 To 200000	4	11
100000 To 200000	5	12
100000 To 200000	6	4
200000 To 300000	1	4
200000 To 300000	2	9
200000 To 300000	3	3
200000 To 300000	4	10
200000 To 300000	5	4
200000 To 300000	6	2
Over 300000	1	6
Over 300000	2	6
Over 300000	3	1
Over 300000	4	1
Over 300000	5	2
Over 300000	6	3

Counts Table

	Price				
City	Up To 100000	100000 To 200000	200000 To 300000	Over 300000	Total
1	5	12	4	6	27
2	18	16	9	6	49
3	5	3	3	1	12
4	5	11	10	1	27
5	6	12	4	2	24
6	2	4	2	3	11
Total	41	58	32	19	150

This report gives the count (frequency) for each unique combination of the table and grouping variables, taken together. In this example, there are no grouping variables.

Example 9 – Pairwise Multiple Comparisons

This example will generate a set of simultaneous, pairwise multiple comparisons of data in Table 1 (page 1271) of Agresti et al. (2008). The data presents the results of a study in which 347 subjects are randomly separated into four groups. These groups received four different treatments (A, B, C, and D). At the end of the study people, the number in each group who had experienced the side-effect of nausea was tabulated. The result of this tabulation is given below.

	Nausea	
Treatment	**No**	**Yes**
A	74	13
B	59	27
C	65	22
D	78	9

A Pearson Chi-Square test on the above counts resulted in a p-value of 0.002. The next question was to find out which groups were different? In order to answer this question, the researchers decided to construct a table of simultaneous confidence intervals for the differences among the four group proportions. Two methods are available to generate these confidence intervals as well as a third method that only gives p-values. We will look at reports using all three methods in this example.

Rather than enter the data on a dataset, it will be entered directly into the template window.

Setup

To run this example, complete the following steps:

1 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 9** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Two-Way Table**
Rows	**4**
Columns	**2**
Two-Way Table of Counts	
Column Title	**Nausea**
Column 1 Label	**No**
Column 2 Label	**Yes**
Row Title	**Treatment**
Row 1 Label	**A**
Row 2 Label	**B**
Row 3 Label	**C**
Row 4 Label	**D**
Column 1 Counts	**74 59 65 78**
Column 2 Counts	**13 27 22 9**

Reports Tab
Show Individual Tables.......................... **Checked**
Counts .... **Checked**
Row Percentages .... **Checked**
Tests for Row-Column Independence .... **Checked**
(Pearson's Chi-Square, Fisher's Exact)
Pearson's Chi-Square Test.................. **Checked**

Multiple Comparisons Tab
Success Category................................. **Second Category**
Pairwise Wald Tests and C.I.'s of Differences .... **Checked**
Count Adjustment **1**
Pairwise Score Tests and C.I.'s of Differences .... **Checked**
Count Adjustment **0**
Score C.I. Accuracy............................. **3 Decimal Places**
Pairwise Tests of Differences using the Angular Transformation .... **Checked**

2 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Treatment	Nausea No	Yes	Total
A	74	13	87
B	59	27	86
C	65	22	87
D	78	9	87
Total	276	71	347

Row Percentages Table

Treatment	Nausea No	Yes	Total
A	85.1%	14.9%	100.0%
B	68.6%	31.4%	100.0%
C	74.7%	25.3%	100.0%
D	89.7%	10.3%	100.0%
Total	79.5%	20.5%	100.0%

Tests for Row-Column Independence
(Treatment by Nausea)
H0: "Treatment" and "Nausea" are independent.
H1: "Treatment" and "Nausea" are associated (not independent).

Test	Type	Chi-Square Value	DF	Prob Level	Reject H0 at α = 0.05?
Pearson's Chi-Square	2-Sided	14.662	3	0.0021	Yes

Pairwise Adjusted Wald Tests and Simultaneous C.I.'s of Proportion Difference
(Treatment by Nausea)
Test Method: Adjusted Wald test
Success Event: Nausea = Yes
Count Adjustment Value: 1

Comparison Groups	Count	Prop	Prop Diff	Student Range Q	P-Value	Lower 95.0% Conf Limit	Upper 95.0% Conf Limit
A	**87**	**0.149**					
- B	86	0.314	-0.165	3.618	0.0515	-0.322	0.001
- C	87	0.253	-0.103	2.369	0.3365	-0.256	0.054
- D	87	0.103	0.046	1.244	0.7966	-0.086	0.176
B	**86**	**0.314**					
- A	87	0.149	0.165	3.618	0.0515	-0.001	0.322
- C	87	0.253	0.061	1.243	0.7968	-0.115	0.234
- D	87	0.103	0.211	4.861	0.0033*	0.052	0.360
C	**87**	**0.253**					
- A	87	0.149	0.103	2.369	0.3365	-0.054	0.256
- B	86	0.314	-0.061	1.243	0.7968	-0.234	0.115
- D	87	0.103	0.149	3.610	0.0522	-0.001	0.293
D	**87**	**0.103**					
- A	87	0.149	-0.046	1.244	0.7966	-0.176	0.086
- B	86	0.314	-0.211	4.861	0.0033*	-0.360	-0.052
- C	87	0.253	-0.149	3.610	0.0522	-0.293	0.001

* The proportion difference is significant at the 0.05 level.

Pairwise Comparison Score Tests and Simultaneous C.I.'s of Proportion Difference
(Treatment by Nausea)
Test Method: Score Test
Success Event: Nausea = Yes
Count Adjustment Value: 0

Comparison Groups	Count	Prop	Prop Diff	Student Range Q	P-Value	Lower 95.0% Conf Limit	Upper 95.0% Conf Limit
A	**87**	**0.149**					
- B	86	0.314	-0.165	3.695	0.0445*	-0.325	0.000
- C	87	0.253	-0.103	2.427	0.3149	-0.260	0.054
- D	87	0.103	0.046	1.293	0.7811	-0.089	0.184
B	**86**	**0.314**					
- A	87	0.149	0.165	3.695	0.0445*	0.000	0.325
- C	87	0.253	0.061	1.263	0.7907	-0.114	0.234
- D	87	0.103	0.211	4.982	0.0024*	0.055	0.364
C	**87**	**0.253**					
- A	87	0.149	0.103	2.427	0.3149	-0.054	0.260
- B	86	0.314	-0.061	1.263	0.7907	-0.234	0.114
- D	87	0.103	0.149	3.714	0.0429*	0.000	0.299
D	**87**	**0.103**					
- A	87	0.149	-0.046	1.293	0.7811	-0.184	0.089
- B	86	0.314	-0.211	4.982	0.0024*	-0.364	-0.055
- C	87	0.253	-0.149	3.714	0.0429*	-0.299	0.000

* The proportion difference is significant at the 0.05 level.

These two reports give p-values and simultaneous confidence intervals of the proportion difference using the adjusted Wald and the Score methods. In most cases the differences between the results of these two methods are minor.

Student Range Q

This is the value of the test statistic.

P-Value

The p-value is the probability of rejecting the null hypothesis that the proportions are equal. This probability is adjusted for the fact that you are conducting several hypothesis tests.

Confidence Limits

These are the lower and upper confidence limits for the simultaneous, two-sided confidence intervals of the difference of the proportions shown on this line of the report.

Pairwise Wald Tests of the Difference in Proportions using the Angular Transformation
(Treatment by Nausea)
Test Method: Wald test with angular (arcsin square root) transformation
Success Event: Nausea = Yes

Comparison Groups	Count	Prop	Prop Diff	Angular Trans Diff	Student Range Q	P-Value
A	**87**	**0.1494**				
- B	86	0.314	-0.165	-11.153	3.631	0.0502
- C	87	0.253	-0.103	-7.318	-2.389	1.0000
- D	87	0.103	0.046	3.880	1.267	0.7896
B	**86**	**0.3140**				
- A	87	0.149	0.165	11.153	3.631	0.0502
- C	87	0.253	0.061	3.835	1.249	0.7952
- D	87	0.103	0.211	15.033	4.894	0.0030*
C	**87**	**0.2529**				
- A	87	0.149	0.103	7.318	2.389	0.3290
- B	86	0.314	-0.061	-3.835	1.249	0.7952
- D	87	0.103	0.149	11.198	3.656	0.0479*
D	**87**	**0.1034**				
- A	87	0.149	-0.046	-3.880	-1.267	1.0000
- B	86	0.314	-0.211	-15.033	4.894	0.0030*
- C	87	0.253	-0.149	-11.198	-3.656	1.0000

* The proportion difference is significant at the 0.05 level.

This report gives the p-values of the proportion difference using the angular transformation (arcsin square-root).

Student Range Q

This is the value of the test statistic.

P-Value

The p-value is the probability of rejecting the null hypothesis that the proportions are equal. This probability is adjusted for the fact that you are conducting several hypothesis tests.

Pairwise Comparison Score Tests and Simultaneous C.I.'s of Odds Ratio
(Treatment by Nausea)

Test Method: Score Test
Success Event: Nausea = Yes
Count Adjustment Value: 0

Comparison Groups	Count	Prop	Odds Ratio	Student Range Q	P-Value	Lower 95.0% Conf Limit	Upper 95.0% Conf Limit
A	**87**	**0.149**					
- B	86	0.314	0.384	3.629	0.0504	0.148	1.000
- C	87	0.253	0.519	2.407	0.3224	0.196	1.380
- D	87	0.103	1.523	1.290	0.7821	0.484	4.769
B	**86**	**0.314**					
- A	87	0.149	2.605	3.629	0.0504	1.000	6.762
- C	87	0.253	1.352	1.261	0.7915	0.572	3.192
- D	87	0.103	3.966	4.823	0.0036*	1.381	11.298
C	**87**	**0.253**					
- A	87	0.149	1.927	2.407	0.3224	0.725	5.107
- B	86	0.314	0.740	1.261	0.7915	0.313	1.748
- D	87	0.103	2.933	3.642	0.0492*	1.003	8.516
D	**87**	**0.103**					
- A	87	0.149	0.657	1.290	0.7821	0.210	2.066
- B	86	0.314	0.252	4.823	0.0036*	0.089	0.724
- C	87	0.253	0.341	3.642	0.0492*	0.117	0.997

* The odds ratio is significant at the 0.05 level.

This report gives the p-values and simultaneous confidence intervals of the odds ratios using the adjusted Wald and the Score methods.

Student Range Q

This is the value of the test statistic.

P-Value

The p-value is the probability of rejecting the null hypothesis that the odds ratio is equal to one. This probability is adjusted for the fact that you are conducting several hypothesis tests.

Confidence Limits

These are the lower and upper confidence limits for the simultaneous, two-sided confidence intervals of the odds ratios of the proportions shown on this line of the report.

Example 10 – Multiple Comparisons versus a Control

This example will generate a set of simultaneous, many-to-one multiple comparisons of data in Table 1 (page 1271) of Agresti et al. (2008). The data presents the results of a study in which 347 subjects are randomly separated into four groups. These groups received four different treatments (A, B, C, and D). In this example, Group D is the control (placebo).

At the end of the study people, the number in each group who had experienced the side-effect of nausea was tabulated. The result of this tabulation is given below.

	Nausea	
Treatment	**No**	**Yes**
A	74	13
B	59	27
C	65	22
D	78	9

A Pearson Chi-Square test on the above counts resulted in a p-value of 0.002. The next question was to find out which treatment groups were different from the control group? In order to answer this question, the researchers decided to construct a table of simultaneous confidence intervals for the differences between the three treatment groups and the control group. Two statistical methods are available to generate these confidence intervals: Dunnett and Bonferroni. We will look at reports using both of these methods in this example. Also, we will look at the one-sided confidence intervals that can be generated with the Dunnett method.

Rather than enter the data on a dataset, it will be entered directly into the template window.

Setup

To run this example, complete the following steps:

1 Specify the Contingency Tables (Crosstabs / Chi-Square Test) procedure options

- Find and open the **Contingency Tables (Crosstabs / Chi-Square Test)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 10** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Type of Data Input	**Two-Way Table**
Rows	**4**
Columns	**2**
Two-Way Table of Counts	
Column Title	**Nausea**
Column 1 Label	**No**
Column 2 Label	**Yes**
Row Title	**Treatment**
Row 1 Label	**A**
Row 2 Label	**B**
Row 3 Label	**C**
Row 4 Label	**D**
Column 1 Counts	**74 59 65 78**
Column 2 Counts	**13 27 22 9**

Reports Tab
Show Individual Tables.......................... **Checked**
Counts .... **Checked**
Row Percentages .... **Checked**
Tests for Row-Column Independence .... **Checked**
(Pearson's Chi-Square, Fisher's Exact)
Pearson's Chi-Square Test.................. **Checked**

Multiple Comparisons Tab
Success Category................................. **Second Category**
Control Group .... **Last Group**
Two-Sided Dunnett Tests and C.I.'s....... **Checked**
versus a Control
Count Adjustment **1**
Two-Sided Bonferroni Tests and C.I.'s... **Checked**
versus a Control
Count Adjustment **1**
Lower One-Sided Dunnett Tests and ..... **Checked**
C.I.'s versus a Control
Count Adjustment **1**
Upper One-Sided Dunnett Tests and ..... **Checked**
C.I.'s versus a Control
Count Adjustment **1**
Dunnett Score Tests and C.I.'s of........... **Checked**
Odds Ratios
Count Adjustment **0**

2 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Counts Table

Treatment	Nausea No	Yes	Total
A	74	13	87
B	59	27	86
C	65	22	87
D	78	9	87
Total	276	71	347

Row Percentages Table

Treatment	Nausea No	Yes	Total
A	85.1%	14.9%	100.0%
B	68.6%	31.4%	100.0%
C	74.7%	25.3%	100.0%
D	89.7%	10.3%	100.0%
Total	79.5%	20.5%	100.0%

Tests for Row-Column Independence
(Treatment by Nausea)
H0: "Treatment" and "Nausea" are independent.
H1: "Treatment" and "Nausea" are associated (not independent).

Test	Type	Chi-Square Value	DF	Prob Level	Reject H0 at α = 0.05?
Pearson's Chi-Square	2-Sided	14.662	3	0.0021	Yes

Two-Sided Dunnett Tests and C.I.'s versus a Control of Proportion Difference
(Treatment by Nausea)
Test Method: Adjusted Wald test using Dunnett's distribution
Success Event: Nausea = Yes
Control Group: D
Count Adjustment Value: 1

Comparison Groups	Count	Prop	Prop - Control	Test Statistic Q	P-Value	Lower 95.0% Conf Limit	Upper 95.0% Conf Limit
D	**87**	**0.103**					
A	87	0.149	0.046	0.880	0.3722	-0.075	0.165
B	86	0.314	0.211	3.437	0.0009*	0.065	0.346
C	87	0.253	0.149	2.553	0.0145*	0.012	0.280

* The proportion difference is significant at the 0.05 level.

Two-Sided Bonferroni Tests and C.I.'s versus a Control of Proportion Difference
(Treatment by Nausea)
Test Method: Adjusted Wald test using Bonferroni adjustment
Success Event: Nausea = Yes
Control Group: D
Count Adjustment Value: 1

Comparison Groups	Count	Prop	Prop - Control	Chi-Square Q	P-Value	Lower 95.0% Conf Limit	Upper 95.0% Conf Limit
D	**87**	**0.103**					
A	87	0.149	0.046	0.880	1.000	-0.077	0.167
B	86	0.314	0.211	3.437	0.002*	0.062	0.349
C	87	0.253	0.149	2.553	0.032*	0.009	0.283

* The proportion difference is significant at the 0.05 level.

Lower One-Sided Dunnett Tests and Simultaneous C.I.'s versus a Control of Proportion Difference
(Treatment by Nausea)
Test Method: Adjusted Wald test using Dunnett's distribution
Success Event: Nausea = Yes
Control Group: D
Count Adjustment Value: 1

Comparison Groups	Count	Prop	Prop - Control	Test Statistic Q	P-Value	Lower 95.0% Conf Limit
D	**87**	**0.103**				
A	87	0.149	0.046	0.880	0.372	-0.060
B	86	0.314	0.211	3.437	0.001*	0.082
C	87	0.253	0.149	2.553	0.015*	0.028

* The proportion difference is significant at the 0.05 level.

Upper One-Sided Dunnett Tests and Simultaneous C.I.'s versus a Control of Proportion Difference (Treatment by Nausea)

Test Method: Adjusted Wald test using Dunnett's distribution
Success Event: Nausea = Yes
Control Group: D
Count Adjustment Value: 1

Comparison Groups	Count	Prop	Prop - Control	Test Statistic Q	P-Value	Upper 95.0% Conf Limit
D	**87**	**0.103**				
A	87	0.149	0.046	0.880	0.372	0.150
B	86	0.314	0.211	3.437	0.001*	0.329
C	87	0.253	0.149	2.553	0.015*	0.264

* The proportion difference is significant at the 0.05 level.

These four reports give p-values and simultaneous confidence intervals of the proportion difference using the adjusted Wald test methods and Dunnett's distribution.

Test Statistic Q

This is the value of the test statistic (Wald ratio).

P-Value

The p-value is the probability of rejecting the null hypothesis that the proportions are equal. This probability is adjusted for the fact that you are conducting several hypothesis tests.

Confidence Limits

These are the confidence limits for the simultaneous confidence interval of the difference in the proportions shown on this line of the report.

Two-Sided Dunnett Tests and Simultaneous C.I.'s versus a Control of Odds Ratio (Treatment by Nausea)

Test Method: Adjusted Score test using Dunnett's distribution
Success Event: Nausea = Yes
Control Group: D
Count Adjustment Value: 0

Comparison Groups	Count	Prop	Odds Ratio	Test Statistic Q	P-Value	Lower 95.0% Conf Limit	Upper 95.0% Conf Limit
D	**87**	**0.103**					
A	87	0.149	1.523	0.912	0.686	0.531	4.354
B	86	0.314	3.966	3.410	0.002*	1.505	10.388
C	87	0.253	2.933	2.576	0.027*	1.096	7.815

* The odds ratio is significant at the 0.05 level.

This report give p-values and simultaneous confidence intervals of the odds ratio using the adjusted score test and Dunnett's distribution.

Test Statistic Q

This is the value of the test statistic (Wald ratio for testing the null hypothesis that the odds ratio = 1).

P-Value

The p-value is the probability of rejecting the null hypothesis that the ratio of the treatment group odds and the control group odds is one. This probability is adjusted for the make that several hypothesis tests are being made.

Confidence Limits

These are the confidence limits for the simultaneous confidence interval of the odds ratio of the group proportion shown on this line with the control group.