

Chapter 9

Table Analysis

Chapter Table of Contents

Introduction	179
Association in a 2×2 Table	181
Exact Test	187
Association in Sets of Tables	193
Observer Agreement	201
References	206

Chapter 9

Table Analysis

Introduction

Often you need to analyze the information in a table, sometimes called a contingency table or a crossclassification table. You may analyze a single table, or you may analyze a set of tables. You are also often concerned with evaluating the presence of *association* in a table, or whether there is some sort of relationship between the variable determining the rows of the table and the variable determining the columns of the table. If there is an inherent ordering in the rows or columns of the table, the association may be linear. Various chi-square statistics such as the Pearson chi-square and the likelihood ratio chi-square are used to assess association.

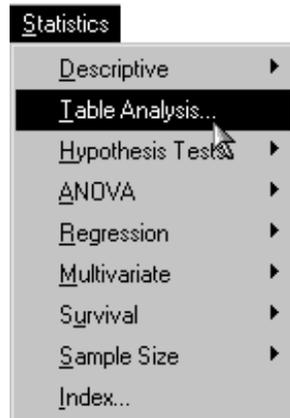


Figure 9.1. Table Analysis Selection Menu

Besides assessing the presence of association, you may also be interested in computing a *measure of association*, or a statistic that provides some understanding of the strength of the association. The odds ratio is a standard measure of association often used in medical and epidemiological studies.

Using the Table Analysis task, not only can you analyze a single table, but you can also analyze sets of tables. This provides a way to control, or adjust for, a covariate, while assessing association of the rows and columns of the tables. Extended Mantel-Haenszel statistics, also called Cochran-Mantel-Haenszel statistics, provide a way to utilize all the information in the constituent tables in a test for the hypothesis of association. Tables may also contain information from observer agreement studies in which the evaluations or assessments of two different observers are collected. Statistics called measures of agreement assess how closely the observers agree.

The Table Analysis task provides chi-square tests of association for the $r \times c$ table, including statistics such as the Pearson chi-square and likelihood ratio test, and it also computes extended Mantel-Haenszel tests for sets of tables. Fisher's exact test can be computed for both the 2×2 and $r \times c$ table. In addition, the Table Analysis task also provides measures of association such as the odds ratio and relative risk for the 2×2 table as well as gamma, tau-*b*, Somer's *D*, and the Pearson and Spearman correlation coefficients. In addition, you can obtain measures of agreement such as the kappa coefficient and the weighted kappa coefficient. McNemar's test is produced for the 2×2 table.

The examples in this chapter demonstrate how you can use the Analyst Application to analyze tables, including assessing the presence of association in a table and sets of tables and assessing observer agreement.

Association in a 2×2 Table

The most basic table is a 2×2 table. Usually, the columns represent some sort of outcome, often yes or no, and the rows represent levels of a factor that may influence the outcome. Suppose, for example, that researchers were investigating the properties of a new “ouch-less” bandaid for children. Interest lies in whether those children trying the test bandaid recorded fewer complaints on removal than those children using a regular bandaid. You can address this question by forming the two-way table of bandaid type and complaint status and then assessing the association between the rows and columns of that table.

Open the Bandaid Data Set

These data are provided as the **Bandaid** data set in the Analyst Sample Library. To open the **Bandaid** data set, follow these steps:

1. Select **Tools** → **Sample Data** . . .
2. Select **Bandaid**.
3. Click **OK** to create the sample data set in your **Sasuser** directory.
4. Select **File** → **Open By SAS Name** . . .
5. Select **Sasuser** from the list of **Libraries**.
6. Select **Bandaid** from the list of members.
7. Click **OK** to bring the **Bandaid** data set into the data table.

Figure 9.2 displays the data table containing these data. Note that the data are in frequency form, with the variable **count** containing the frequencies of the profile contained in each row of the table. The variable **type** is the type of bandaid tested and the variable **outcome** is the status of complaints.

Bandaid (Browse)			
	type	outcome	count
1	regular	complain	14
2	regular	no	16
3	test	complain	10
4	test	no	30

Figure 9.2. Data Set Bandaid in the Data Table

Specify the Table

To construct the appropriate two-way table and request tests of association, follow these steps:

1. Select **Statistics** → **Table Analysis** . . .
2. Select **type** from the candidate list as the **Row** variable.
3. Select **outcome** from the candidate list as the **Column** variable.
4. Select **count** from the candidate list as the **Cell Counts** variable.

Figure 9.3 displays the resulting dialog.

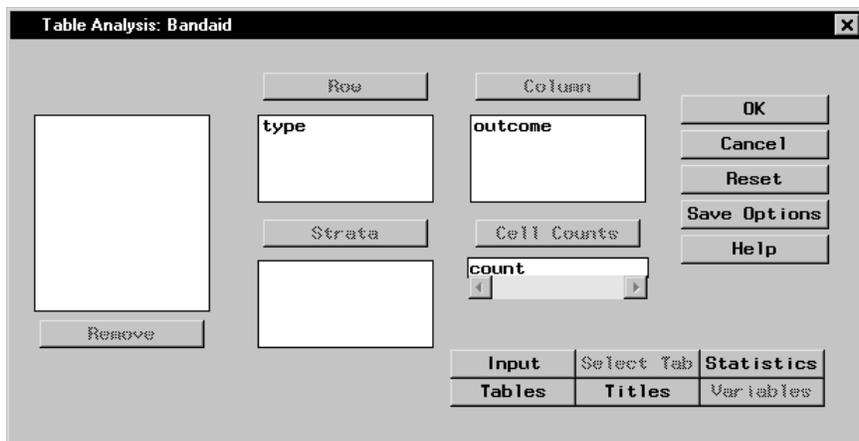


Figure 9.3. Table Analysis Task for Bandid Study

Request Tests and Measures of Association

By selecting the rows and columns of the table, you have requested the construction of a 2×2 table. To request chi-square tests of association and the odds ratio, which is a measure of association, follow these steps:

1. Click on the **Statistics** button.
2. Select **Chi-square statistics**.
3. Select **Measures of association**.
4. Click **OK**.

Figure 9.4 displays the Statistics dialog.

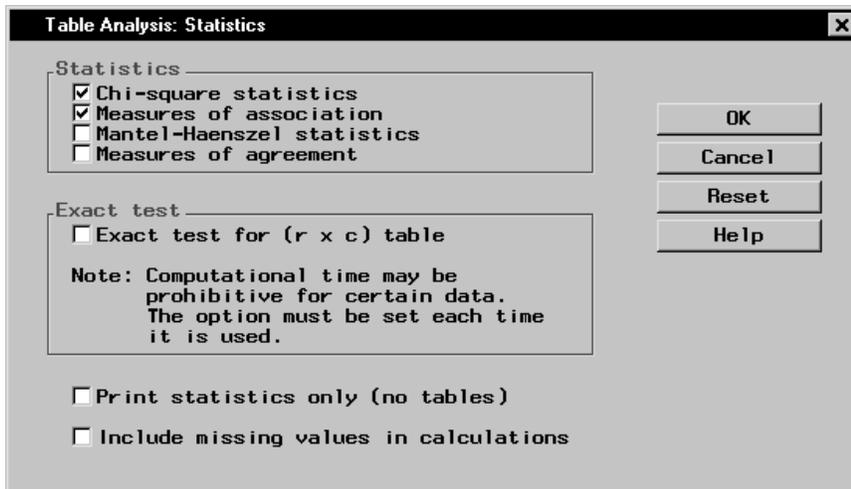


Figure 9.4. Statistics Dialog

Finally, in order to customize the form of the displayed table, follow these steps:

1. Click on the **Tables** button.
2. Select **Observed** under **Frequencies**.
3. Select **Row** under **Percentages**.
4. Click **OK**.

Figure 9.5 displays the resulting Tables dialog.

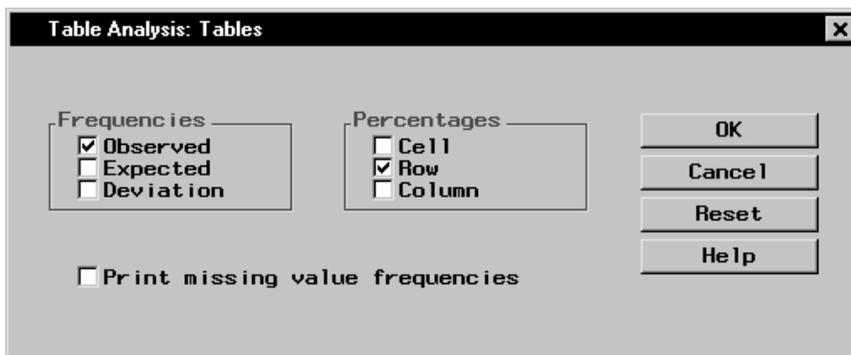


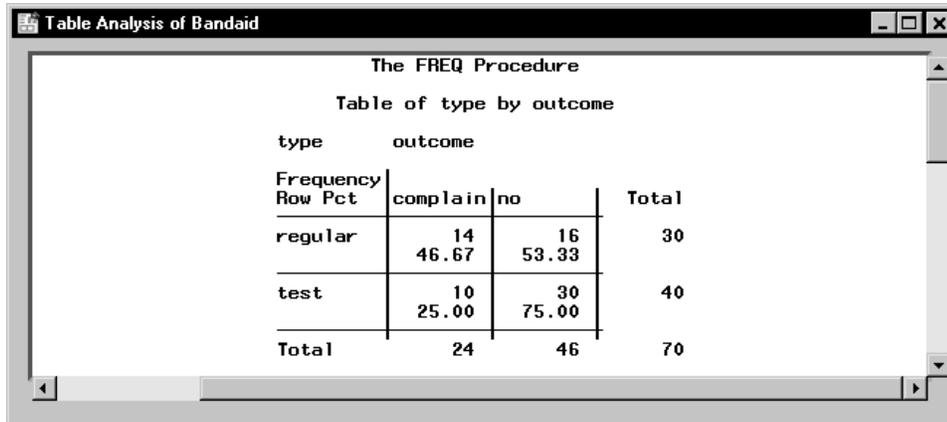
Figure 9.5. Tables Dialog

This requests that only the raw frequencies and the row percentages be listed in the printed table cell.

Click **OK** in the Table Analysis dialog to perform the analysis.

Review the Results

The frequency table is displayed in Figure 9.6. Note that 46 percent of those children getting regular bandaids had complaints about irritation when their bandaid was removed, compared to 25 percent of those children receiving the test bandaid.



The screenshot shows a window titled "Table Analysis of Bandaid". Inside, it displays "The FREQ Procedure" and a "Table of type by outcome". The table is as follows:

type	outcome		Total
	complain	no	
regular	14 46.67	16 53.33	30
test	10 25.00	30 75.00	40
Total	24	46	70

Figure 9.6. Frequency Table for Bandaid Data

Figure 9.7 contains the table of computed chi-square statistics for this table. The Pearson chi-square statistic, labeled “Chi-Square,” has a value of 3.57 and an associated p -value of 0.0588 with 1 degree of freedom. If you were doing strict hypothesis testing, you would not reject the hypothesis of no association at the $\alpha = 0.05$ level of significance. However, researchers in this case found enough evidence in this pilot study to continue looking into the new product.

Table Analysis of Bandid

Statistics for Table of type by outcome

Statistic	DF	Value	Prob
Chi-Square	1	3.5719	0.0588
Likelihood Ratio Chi-Square	1	3.5655	0.0590
Continuity Adj. Chi-Square	1	2.6749	0.1019
Mantel-Haenszel Chi-Square	1	3.5208	0.0606
Phi Coefficient		0.2259	
Contingency Coefficient		0.2203	
Cramer's V		0.2259	

Fisher's Exact Test

Cell (1,1) Frequency (F)	14
Left-sided Pr <= F	0.9840
Right-sided Pr >= F	0.0512
Table Probability (P)	0.0351
Two-sided Pr <= P	0.0769

Figure 9.7. Chi-Square Statistics for Bandid Data

Several other chi-square statistics also appear in this output, such as the likelihood ratio chi-square and the Mantel-Haenszel chi-square. These statistics are asymptotically equivalent.

Table Analysis of Bandid

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	2.6250	0.9530	7.2307
Cohort (Co11 Risk)	1.8667	0.9656	3.6085
Cohort (Co12 Risk)	0.7111	0.4865	1.0394

Sample Size = 70

Figure 9.8. Odds Ratio for Bandid Data

Figure 9.8 contains the table of relative risk estimates including the odds ratio, which is labeled “Case-Control.” The odds ratio is the ratio of the odds of having an outcome for one group versus another.

When the odds ratio has the value 1, you have equal odds of having the outcome. When the odds ratio is greater than 1, one group has greater odds of an outcome than the other.

The odds ratio has a value of 2.62, which means that the odds of a complaint are 2.62 times higher for those children using the regular bandaid than for those using the test bandaid.

Exact Test

You may have noticed that the preceding statistical output also included a test called Fisher's Exact test. When the sample size for the test of association of a table does not meet the usual guidelines (generally 20-25 total observations for a 2×2 table, with 80 percent of the table cells having counts greater than 5), an exact test may be a useful strategy.

The following data illustrate where an exact test may be appropriate. A marketing research firm took a sample of members at a health club and asked them a series of questions. They were interested in gathering information that could help their clients decide on audiences to target for new magazines. One of the questions was what activity the member considered his or her primary activity at the club. Another question was whether the member was considering making a major diet change. The researchers were interested in what types of sports magazines in which to place ads for a new food and nutrition magazine.

Open the Gym Data Set

These data are provided as the **Gym** data set in the Analyst Sample Library. To open the **Gym** data set, follow these steps:

1. Select **Tools** → **Sample Data** . . .
2. Select **Gym**.
3. Click **OK** to create the sample data set in your **Sasuser** directory.
4. Select **File** → **Open By SAS Name** . . .
5. Select **Sasuser** from the list of **Libraries**.

6. Select **Gym** from the list of members.
7. Click **OK** to bring the **Gym** data set into the data table.

Figure 9.9 displays the data table containing these data. Note that the data are in frequency form, with the variable **count** containing the frequencies of the profile contained in each row of the table. The variable **activity** contains the type of activity, which can be aerobics, yoga, weightlifting, team sports such as volleyball and basketball leagues, and cross-training. The variable **DietChange** indicates whether the member was contemplating a change in diet.

Gym (Browse)			
	activity	DietChange	count
1	aerobics	yes	13
2	aerobics	no	8
3	yoga	yes	3
4	yoga	no	2
5	weights	yes	3
6	weights	no	19
7	team	yes	12
8	team	no	16
9	cross	yes	11
10	cross	no	13

Figure 9.9. Data Set Gym in the Data Table

Specify the Table

To construct the appropriate two-way table and request tests of association, follow these steps:

1. Select **Statistics** → **Table Analysis** . . .
2. Select **activity** from the candidate list as the **Row** variable.
3. Select **DietChange** from the candidate list as the **Column** variable.
4. Select **count** from the candidate list as the **Cell Counts** variable.

Figure 9.10 displays the resulting dialog.

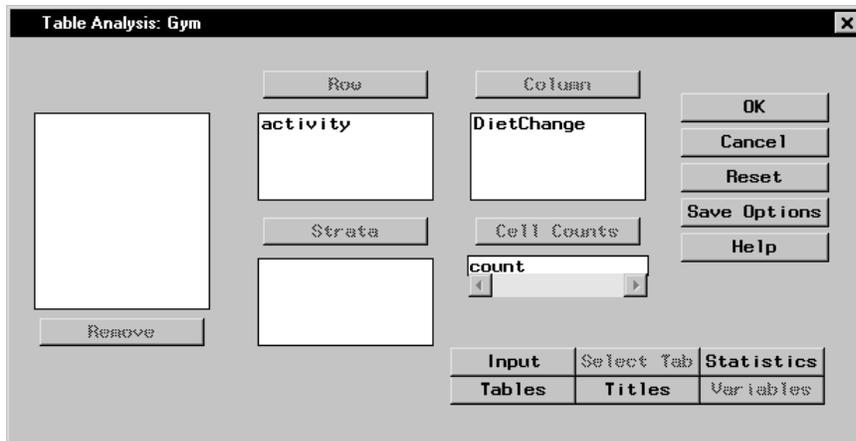


Figure 9.10. Table Analysis Task for Health Club Study

Request Tests and Measures of Association

By selecting the rows and columns of the table, you have requested the construction of a 5×2 table. To request chi-square tests of association, follow these steps:

1. Click on the **Statistics** button.
2. Select **Chi-square statistics**.
3. Click **OK**.

Note that the Tables dialog specifications (see Figure 9.5) made in the previous analysis remain in effect. Therefore, both frequencies and row percentages are produced for this analysis.

Click **OK** in the Table Analysis dialog to perform the analysis.

Review the Results

The frequency table is displayed in Figure 9.11. Note that 62 percent of those members participating in aerobics were considering a diet change and so were 60 percent of yoga practitioners. Eighty-six percent of those members lifting weights were not considering a diet change. Of those members playing a team sport or who considered themselves cross-trainers, the majority of members were not considering a diet change, but not by a wide margin.

The screenshot shows a window titled "Table Analysis of Gym" containing the output of the FREQ procedure. The output is a table titled "Table of activity by DietChange". The table has columns for "activity", "DietChange" (with sub-columns "no" and "yes"), and "Total". The rows represent different activities: aerobics, cross, team, weights, and yoga. Each cell contains the frequency count and the row percentage.

activity	DietChange		Total
	no	yes	
aerobics	8 38.10	13 61.90	21
cross	13 54.17	11 45.83	24
team	16 57.14	12 42.86	28
weights	19 86.36	3 13.64	22
yoga	2 40.00	3 60.00	5
Total	58	42	100

Figure 9.11. Frequency Table for Gym Data

Figure 9.12 contains the table of chi-square statistics computed for this table. The Pearson chi-square statistic has a value of 11.4993 and an associated p -value of 0.0215 with 4 degrees of freedom. If you were doing strict hypothesis testing, you would reject the hypothesis of no association at the $\alpha = 0.05$ level of significance. However, if you look at Figure 9.11, you see that three table cells have a count of less than 5, which violates one of the sample size guidelines for the asymptotic tests. Thus, you may want to compute the exact test for these data.

The screenshot shows a window titled "Table Analysis of Gym". Inside, there is a table of statistics for a table of activity by DietChange. The table has four columns: Statistic, DF, Value, and Prob. Below the table, it indicates a Sample Size of 100.

Statistic	DF	Value	Prob
Chi-Square	4	11.4993	0.0215
Likelihood Ratio Chi-Square	4	12.5455	0.0137
Mantel-Haenszel Chi-Square	1	5.5259	0.0187
Phi Coefficient		0.3391	
Contingency Coefficient		0.3211	
Cramer's V		0.3391	

Sample Size = 100

Figure 9.12. Chi-square Statistics for Gym Data

Request the Exact Test

To request the exact test, simply return to the Table Analysis task and open the Statistics dialog. All of the settings you have previously selected for the table analysis are still in place. You need only request the additional exact test.

1. Select **Statistics** → **Table Analysis . . .**
2. Click on the **Statistics** button.
3. Select **Exact test for (r × c) table**.
4. Click **OK**.
5. Click **OK** in the main Table Analysis dialog to perform the analysis.

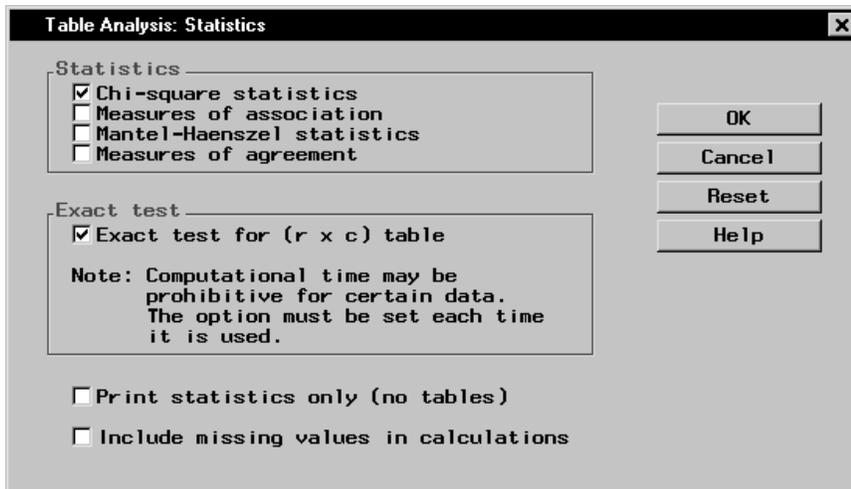


Figure 9.13. Statistics Dialog

Figure 9.13 displays the resulting dialog. Notice the warning that exact test computations may take an excessive amount of time. This would not be the case with very small cell counts, but it is an issue for other tables.

Review the Results

Figure 9.14 contains the results of this analysis, including the exact test results.

Statistic	DF	Value	Prob
Chi-Square	4	11.4993	0.0215
Likelihood Ratio Chi-Square	4	12.5455	0.0137
Mantel-Haenszel Chi-Square	1	5.5259	0.0187
Phi Coefficient		0.3391	
Contingency Coefficient		0.3211	
Cramer's V		0.3391	

Fisher's Exact Test	
Table Probability (P)	8.421E-06
Pr <= P	0.0139

Sample Size = 100

Figure 9.14. Exact Test Results

The exact test computes a p -value of 0.0139; thus, this test also results in the rejection of the hypothesis of no association in this table. There is some kind of association between the rows of the table and the columns of the table; type of primary activity made a difference in whether members were considering diet changes. Not only does degree of association seem to vary, but so does the direction. The market research company may end up suggesting that sports and fitness magazines be targeted in different ways for the new food and diet magazine ad campaign.

Association in Sets of Tables

After the pilot study on the new ouchless bandaids, the investigators decided to continue their research by conducting a clinical trial in which children at five clinics were tested with the test and regular bandaids. Instead of a single table, the clinical trial produces five tables. In order to assess whether the test bandaids produced fewer complaints than the regular bandaids, you need to assess the association in sets of tables instead of the association in a single table.

Extended Mantel-Haenszel statistics, also known as Cochran-Mantel-Haenszel statistics, provide a way of assessing association between two variables that determine a table while controlling for, or adjusting for, the variables that determine the sets of tables. These variables are also known as stratification variables. In this instance, the statistics can provide a way to assess the association between bandaid type and complaint status while controlling for clinic.

In the first section, the odds ratio was presented as a measure of association. You can also compute an overall odds ratio for a set of tables that has been adjusted for the stratification variables.

The `Studybandaid` data set contains the information collected in this clinical trial and includes data that constitute tables for each of the five clinics.

Open the Studybandaid Data Set

These data are provided as the `Studybandaid` data set in the Analyst Sample Library. To open the `Studybandaid` data set, follow these steps:

1. Select **Tools** → **Sample Data** . . .
2. Select `Studybandaid`.
3. Click **OK** to create the sample data set in your `Sasuser` directory.
4. Select **File** → **Open By SAS Name** . . .
5. Select `Sasuser` from the list of **Libraries**.
6. Select `Studybandaid` from the list of members.
7. Click **OK** to bring the `Studybandaid` data set into the data table.

Figure 9.15 displays the data table containing these data. Note that the data are in frequency form, with the variable `COUNT` containing the frequencies of the profile contained in each row of the table. The column corresponding to the variable `CLINIC` contains the values for the five clinics.

Studybandaid (Browse)				
	clinic	type	outcome	count
1	A	regular	complain	14
2	A	regular	no	17
3	A	test	complain	11
4	A	test	no	31
5	B	regular	complain	22
6	B	regular	no	21
7	B	test	complain	10
8	B	test	no	40
9	C	regular	complain	22
10	C	regular	no	28
11	C	test	complain	15
12	C	test	no	30
13	D	regular	complain	15
14	D	regular	no	18
15	D	test	complain	8
16	D	test	no	29
17	E	regular	complain	20
18	E	regular	no	30
19	E	test	complain	15
20	E	test	no	29

Figure 9.15. Data Set Studybandaid in the Data Table

Specify the Tables

To request individual table tests of association as well as the CMH tests for the association of type of bandaid with complaint outcome, first specify the tables under study.

1. Select **Statistics** → **Table Analysis** . . .
2. Select **type** from the candidate list as the **Row** variable.
3. Select **outcome** from the candidate list as the **Column** variable.
4. Select **clinic** from the candidate list as the **Strata** variable.
5. Select **count** from the candidate list as the **Cell Counts** variable.

Figure 9.16 displays the resulting dialog.

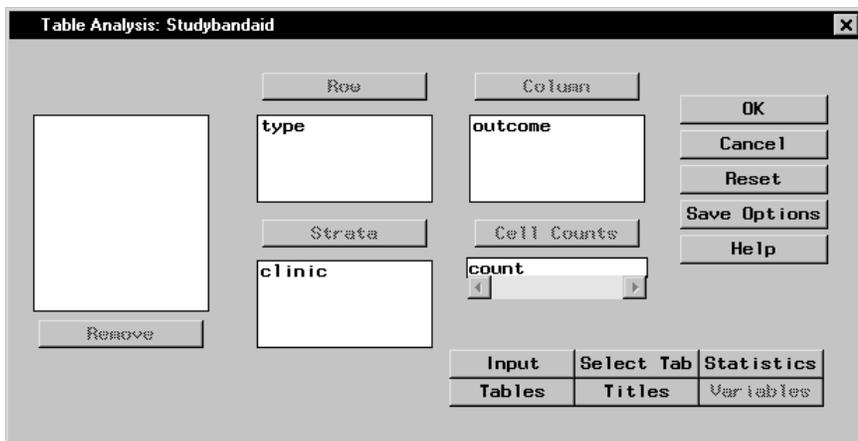


Figure 9.16. Table Analysis Task for Bandaid Study

Request Tests and Measures of Association

Use the Statistics dialog to specify the tests.

1. Click on the **Statistics** button.
2. Select **Chi-square statistics**.
3. Select **Mantel-Haenszel Statistics**.
4. Click **OK**.

Note that the Tables dialog specifications (see Figure 9.5) made previously remain in effect. Therefore, both frequencies and row percentages are produced for this analysis.

Click **OK** in the Table Analysis dialog to perform the analysis.

Review the Results

The results produced include individual tables, individual table statistics, and the summary chi-square statistics.

Figure 9.17 contains the frequency table for clinic A. Figure 9.18 contains the table statistics for clinic A. The Pearson chi-square statistic has the value 2.8505 and a *p*-value of 0.091 with 1 degree of freedom.

Table Analysis of Studybandaid

Table 1 of type by outcome
Controlling for clinic=A

type	outcome		Total
Frequency Row Pct	complain	no	
regular	14 45.16	17 54.84	31
test	11 26.19	31 73.81	42
Total	25	48	73

Figure 9.17. Frequency Table for Clinic A

Table Analysis of Studybandaid

Statistics for Table 1 of type by outcome
Controlling for clinic=A

Statistic	DF	Value	Prob
Chi-Square	1	2.8505	0.0913
Likelihood Ratio Chi-Square	1	2.8403	0.0919
Continuity Adj. Chi-Square	1	2.0703	0.1502
Mantel-Haenszel Chi-Square	1	2.8115	0.0936
Phi Coefficient		0.1976	
Contingency Coefficient		0.1939	
Cramer's V		0.1976	

Fisher's Exact Test

Cell (1,1) Frequency (F)	14
Left-sided Pr <= F	0.9736
Right-sided Pr >= F	0.0753
Table Probability (P)	0.0489
Two-sided Pr <= P	0.1342

Sample Size = 73

Figure 9.18. Table Statistics for Clinic A

Figure 9.19 contains the frequency table for clinic B. Figure 9.20 contains the associated table statistics. The Pearson chi-square statistic has a value of 9.9475 and a corresponding *p*-value of 0.0016.

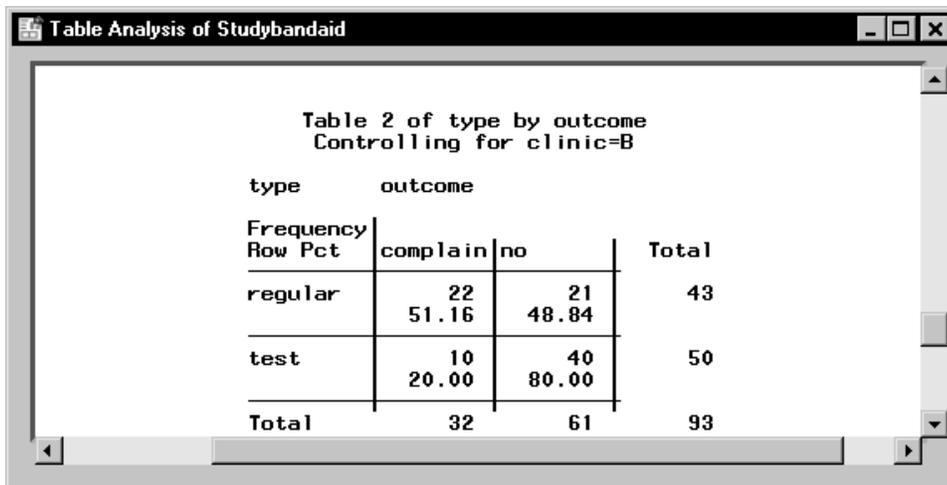


Table Analysis of Studybandaid

Table 2 of type by outcome
Controlling for clinic=B

type	outcome		Total
Frequency Row Pct	complain	no	
regular	22 51.16	21 48.84	43
test	10 20.00	40 80.00	50
Total	32	61	93

Figure 9.19. Frequency Table for Clinic B

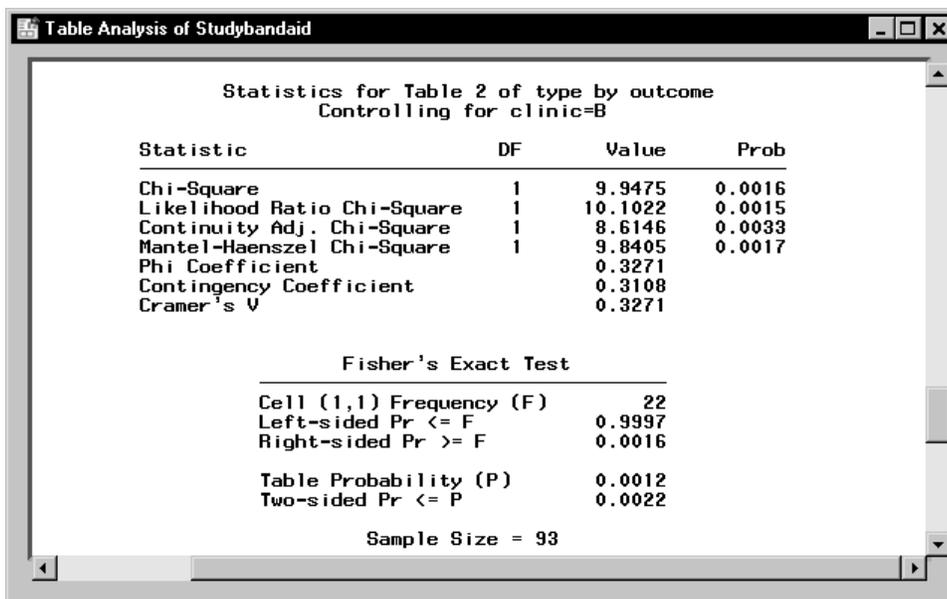


Table Analysis of Studybandaid

Statistics for Table 2 of type by outcome
Controlling for clinic=B

Statistic	DF	Value	Prob
Chi-Square	1	9.9475	0.0016
Likelihood Ratio Chi-Square	1	10.1022	0.0015
Continuity Adj. Chi-Square	1	8.6146	0.0033
Mantel-Haenszel Chi-Square	1	9.8405	0.0017
Phi Coefficient		0.3271	
Contingency Coefficient		0.3108	
Cramer's ψ		0.3271	

Fisher's Exact Test

Cell (1,1) Frequency (F)	22
Left-sided Pr \leq F	0.9997
Right-sided Pr \geq F	0.0016
Table Probability (P)	0.0012
Two-sided Pr \leq P	0.0022

Sample Size = 93

Figure 9.20. Table Statistics for Clinic B

The other individual tables, not printed here, show varying degrees of evidence of association. Clinic C and clinic E appear to have no evidence of association, while clinic D does appear to show evidence of association.

Figure 9.21 displays the results of the CMH analysis.

Summary Statistics for type by outcome Controlling for clinic				
Cochran-Mantel-Haenszel Statistics (Based on Table Scores)				
Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	14.2206	0.0002
2	Row Mean Scores Differ	1	14.2206	0.0002
3	General Association	1	14.2206	0.0002

Figure 9.21. CMH Summary Table

Three versions of the CMH statistic are printed; all have the value 14.2206 and a p -value of 0.0002 with 1 degree of freedom. Your choice of statistic depends on the scale of variables that determine the rows and columns. The General Association statistic always applies. If the columns can be considered ordered, or ordinal, then the Row Mean Score statistic is appropriate as well and is directed at location shifts. If both the columns and rows are ordered, then the Correlation statistic is also appropriate and is directed at linear association. The degrees of freedom of these statistics vary. For more details, refer to Stokes, Davis, and Koch (1995). Note that the sample size requirement for the CMH statistics is that the total (tables combined) sample size be adequate.

In the case of the 2×2 table, all of these statistics are equivalent. Here, you can conclude that type of bandaid is significantly associated with complaint status, controlling for clinic. Figure 9.22 displays the overall relative risk and odds ratios and their confidence bounds.

The screenshot shows a window titled "Table Analysis of Studybandaid". The main content is a table titled "Estimates of the Common Relative Risk (Row1/Row2)". The table has five columns: "Type of Study", "Method", "Value", "95% Confidence Limits", and "Value". The data is as follows:

Type of Study	Method	Value	95% Confidence Limits	Value
Case-Control (Odds Ratio)	Mantel-Haenszel	2.1597	1.4420	3.2348
	Logit	2.1561	1.4331	3.2439
Cohort (Col1 Risk)	Mantel-Haenszel	1.6446	1.2637	2.1402
	Logit	1.6112	1.2355	2.1013
Cohort (Col2 Risk)	Mantel-Haenszel	0.7563	0.6510	0.8787
	Logit	0.7606	0.6545	0.8838

Below the table, the following test results are shown:

Breslow-Day Test for Homogeneity of the Odds Ratios	
Chi-Square	4.4750
DF	4
Pr > ChiSq	0.3455

At the bottom, it states: "Total Sample Size = 425".

Figure 9.22. Odds Ratio

The odds ratio for this study has the value 2.1597 with a confidence bound of (1.4420, 3.2348). This means that those children with the regular bandaid are twice as likely to have complaints as those with the test bandaid or, conversely, that those children with the test bandaid are half as likely to have complaints as those children with the regular bandaid. Since the 95 percent confidence bounds don't include the value 1, this odds ratio is considered to be significantly different from 1.

Note that another test called the Breslow-Day test for Homogeneity of Odds Ratio is also printed. Since the test has a p -value of 0.3455, you would conclude that the hypothesis is not rejected. The sample size requirement for this test is that each individual table has to have sufficient sample size unlike the sample size requirement for the CMH statistics. In this case, since all tables have totals greater than 25, this condition is met.

Observer Agreement

Often, the data represented by a contingency table represents information collected in a study on observer agreement. There may be interest in gathering information on observer error, and such a study may be done as part of testing new processes, training, or tools. Sometimes different observers are studied, and sometimes the same observer is studied at different times or under different conditions.

The members of a northeastern music association were revising their system of conducting local and state-wide high school piano competitions. Instead of using local musicians as judges, they wanted to see if they could proceed more fairly by using one of two trained judges in conjunction with local judges, with whom they needed to come to consensus. In order to see how closely the trained judges match, they did an observer agreement study using some college music students after a training session. Twenty students played one of their current pieces, and both judges rated the performance as good, skilled, or superior.

In order to analyze such data, you form the table with the ratings of one rater forming the rows of the table and the ratings of the other rater forming the columns of the table. The cells of the table are the number of students who fell into the profiles composed of the combination of both ratings. Since there are 3 outcomes, there are 9 possible combinations as represented by the cells of a two-way table. Statistics called measures of agreement are then used to assess the degree of agreement.

Open the Piano Data Set

The Piano data set contains the variables `Rater1` and `Rater2` as well as a frequency variable `count`. These data are provided as the Piano data set in the Analyst Sample Library. To open the Piano data set, follow these steps:

1. Select **Tools** → **Sample Data** . . .
2. Select **Piano**.
3. Click **OK** to create the sample data set in your `Sasuser` directory.

4. Select **File** → **Open By SAS Name . . .**
5. Select **Sasuser** from the list of **Libraries**.
6. Select **Piano** from the list of members.
7. Click **OK** to bring the Piano data set into the data table.

Figure 9.23 displays the data table containing these data. Note that the data are in frequency form, with the variable **count** containing the frequencies of the profile contained in each row of the table. The variable **Rater1** contains the first rater's evaluations and the variable **Rater2** contains the second rater's evaluations.

Piano (Browse)			
	Rater1	Rater2	count
1	good	good	5
2	good	skilled	1
3	good	superior	0
4	skilled	good	2
5	skilled	skilled	5
6	skilled	superior	2
7	superior	good	1
8	superior	skilled	1
9	superior	superior	3

Figure 9.23. Data Set Piano in the Data Table

Specify the Table

To construct the appropriate two-way table, follow these steps:

1. Select **Statistics** → **Table Analysis . . .**
2. Select **Rater1** from the candidate list as the **Row** variable.
3. Select **Rater2** from the candidate list as the **Column** variable.
4. Select **count** from the candidate list as the **Cell Counts** variable.

Figure 9.24 displays the resulting dialog.

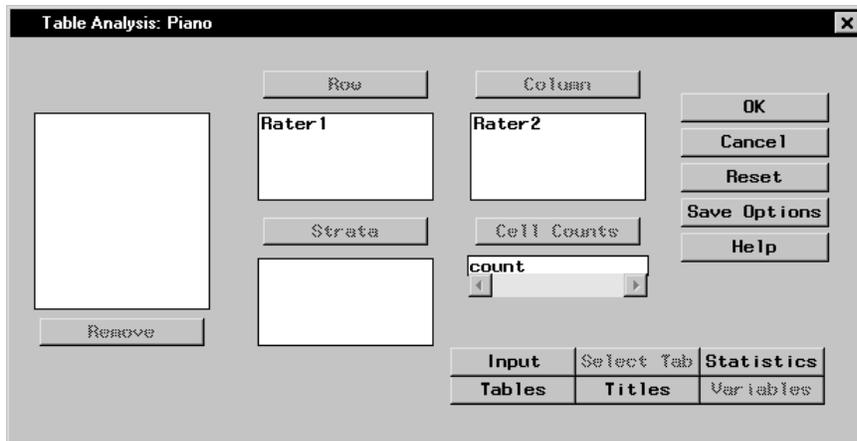


Figure 9.24. Table Analysis Task for Music Study

Request Measures of Agreement

To request measures of agreement, follow these steps:

1. Click on the **Statistics** button.
2. Select **Measures of agreement**.
3. Click **OK**.

Figure 9.25 displays the resulting Statistics dialog. Note that the chi-square tests of association and the measures of association are not appropriate for this type of table.

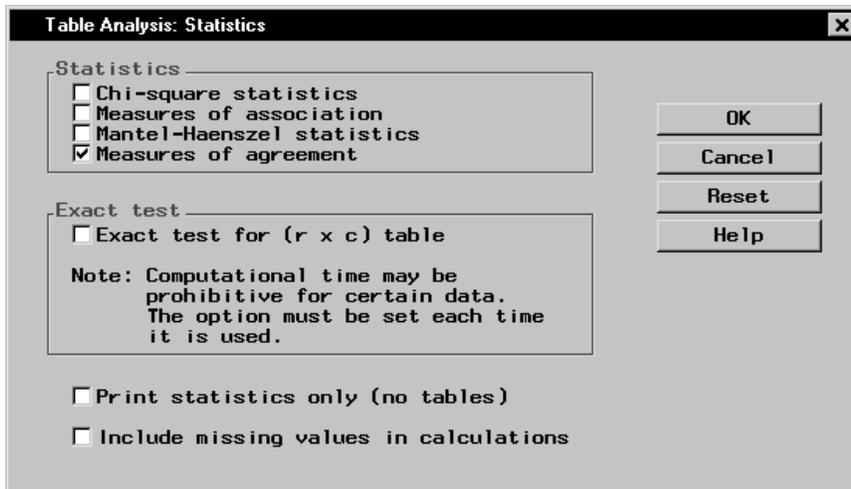


Figure 9.25. Statistics Dialog

Note that the Tables dialog specifications (see Figure 9.5) made previously remain in effect. Therefore, both frequencies and row percentages are produced for this analysis.

Click **OK** in the Table Analysis dialog to perform the analysis.

Review the Results

The frequency table is displayed in Figure 9.26. Note that most of the frequencies occur on the diagonals, which is what you would expect if there is any degree of agreement. However, there are several off-diagonal elements that represent nonagreement. In particular, there is one case of a student rated ‘good’ by Rater2 and ‘superior’ by Rater1. This might be unexpected.

The screenshot shows a window titled "Table Analysis of Piano" with the following content:

The FREQ Procedure

Table of Rater1 by Rater2

Rater1	Rater2			Total
Frequency Row Pct	good	skilled	superior	
good	5 83.33	1 16.67	0 0.00	6
skilled	2 22.22	5 55.56	2 22.22	9
superior	1 20.00	1 20.00	3 60.00	5
Total	8	7	5	20

Figure 9.26. Piano Agreement Frequency Table

Figure 9.27 contains the results for the measures of agreement.

The screenshot shows a window titled "Table Analysis of Piano" with the following content:

Statistics for Table of Rater1 by Rater2

Test of Symmetry

Statistic (S)	1.6667
DF	3
Pr > S	0.6444

Kappa Statistics

Statistic	Value	ASE	95% Confidence Limits	
Simple Kappa	0.4697	0.1597	0.1566	0.7828
Weighted Kappa	0.5210	0.1563	0.2147	0.8272

Sample Size = 20

Figure 9.27. Measures of Agreement

The simple kappa coefficient has a value of 0.4697, with a 95 percent confidence bounds of (0.1566, 0.7828). This suggests modest agreement of ratings. Note that the Bowker's test of symmetry is also printed; this is a test that the probabilities represented by a square table satisfy symmetry.

When you have a 2×2 table, the measure of agreement produced is McNemar's test.

References

SAS Institute Inc. (1999), *SAS/STAT User's Guide, Version 7-1*, Cary, NC: SAS Institute Inc.

Stokes, Maura E., Davis, Charles S., and Koch, Gary G. (1995), *Categorical Data Analysis Using the SAS System*, Cary, NC: SAS Institute Inc.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *The Analyst Application, First Edition*, Cary, NC: SAS Institute Inc., 1999. 476 pp.

The Analyst Application, First Edition

Copyright © 1999 SAS Institute Inc., Cary, NC, USA.

ISBN 1-58025-446-2

All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, by any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute, Inc.

U.S. Government Restricted Rights Notice. Use, duplication, or disclosure of the software by the government is subject to restrictions as set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999

SAS[®] and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. [®] indicates USA registration.

IBM[®], ACF/VTAM[®], AIX[®], APPN[®], MVS/ESA[®], OS/2[®], OS/390[®], VM/ESA[®], and VTAM[®] are registered trademarks or trademarks of International Business Machines Corporation. [®] indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

The Institute is a private company devoted to the support and further development of its software and related services.