SAS/STAT® 9.3 User's Guide

The FREQ Procedure

(Chapter)

SAS® Documentation

This document is an individual chapter from *SAS/STAT® 9.3 User's Guide*.

The correct bibliographic citation for the complete manual is as follows: SAS Institute Inc. 2011. *SAS/STAT® 9.3 User's Guide*. Cary, NC: SAS Institute Inc.

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

1st electronic book, July 2011

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at **support.sas.com/publishing** or call 1-800-727-3228.

Chapter 36

The FREQ Procedure

Contents

Overview: FREQ Procedure

The FREQ procedure produces one-way to *n*-way frequency and contingency (crosstabulation) tables. For two-way tables, PROC FREQ computes tests and measures of association. For *n*-way tables, PROC FREQ provides stratified analysis by computing statistics across, as well as within, strata.

For one-way frequency tables, PROC FREQ computes goodness-of-fit tests for equal proportions or specified null proportions. For one-way tables, PROC FREQ also provides confidence limits and tests for binomial proportions, including tests for noninferiority and equivalence.

For contingency tables, PROC FREQ can compute various statistics to examine the relationships between two classification variables. For some pairs of variables, you might want to examine the existence or strength of any association between the variables. To determine if an association exists, chi-square tests are computed. To estimate the strength of an association, PROC FREQ computes measures of association that tend to be close to zero when there is no association and close to the maximum (or minimum) value when there is perfect association. The statistics for contingency tables include the following:

- chi-square tests and measures
- measures of association
- risks (binomial proportions) and risk differences for 2×2 tables
- odds ratios and relative risks for 2×2 tables
- tests for trend
- tests and measures of agreement
- Cochran-Mantel-Haenszel statistics

PROC FREQ computes asymptotic standard errors, confidence intervals, and tests for measures of association and measures of agreement. Exact *p*-values and confidence intervals are available for many test statistics and measures. PROC FREQ also performs analyses that adjust for any stratification variables by computing statistics across, as well as within, strata for *n*-way tables. These statistics include Cochran-Mantel-Haenszel statistics and measures of agreement.

In choosing measures of association to use in analyzing a two-way table, you should consider the study design (which indicates whether the row and column variables are dependent or independent), the measurement scale of the variables (nominal, ordinal, or interval), the type of association that each measure is designed to detect, and any assumptions required for valid interpretation of a measure. You should exercise care in selecting measures that are appropriate for your data.

Similar comments apply to the choice and interpretation of test statistics. For example, the Mantel-Haenszel chi-square statistic requires an ordinal scale for both variables and is designed to detect a linear association. The Pearson chi-square, on the other hand, is appropriate for all variables and can detect any kind of association, but it is less powerful for detecting a linear association because its power is dispersed over a greater number of degrees of freedom (except for 2×2 tables).

For more information about selecting the appropriate statistical analyses, see Agresti (2007) or Stokes, Davis, and Koch (2000).

Several SAS procedures produce frequency counts; only PROC FREQ computes chi-square tests for one-way to *n*-way tables and measures of association and agreement for contingency tables. Other procedures to consider for counting include the TABULATE and UNIVARIATE procedures. When you want to produce contingency tables and tests of association for sample survey data, use PROC SURVEYFREQ. See Chapter 14, "Introduction to Survey Procedures," for more information. When you want to fit models to categorical data, use a procedure such as CATMOD, GENMOD, GLIMMIX, LOGISTIC, PROBIT, or SURVEYLOGISTIC. See Chapter 8, "Introduction to Categorical Data Analysis Procedures," for more information.

PROC FREQ uses the Output Delivery System (ODS), a SAS subsystem that provides capabilities for displaying and controlling the output from SAS procedures. ODS enables you to convert any of the output from PROC FREQ into a SAS data set. See the section "ODS Table Names" on page 2382 for more information.

PROC FREQ uses ODS Graphics to create graphs as part of its output. For general information about ODS Graphics, see Chapter 21, "Statistical Graphics Using ODS." For specific information about the statistical graphics available with the FREQ procedure, see the PLOTS= option in the TABLES statement and the section "ODS Graphics" on page 2386.

Getting Started: FREQ Procedure

Frequency Tables and Statistics

The FREQ procedure provides easy access to statistics for testing for association in a crosstabulation table.

In this example, high school students applied for courses in a summer enrichment program; these courses included journalism, art history, statistics, graphic arts, and computer programming. The students accepted were randomly assigned to classes with and without internships in local companies. Table 36.1 contains counts of the students who enrolled in the summer program by gender and whether they were assigned an internship slot.

Table 36.1 Summer Enrichment Data

		Enrollment		
Gender	**Internship**	**Yes**	**No**	**Total**
boys	yes	35	29	64
boys	no	14	27	41
girls	yes	32	10	42
girls	no	53	23	76

The SAS data set SummerSchool is created by inputting the summer enrichment data as cell count data, or providing the frequency count for each combination of variable values. The following DATA step statements create the SAS data set SummerSchool:

```
data SummerSchool;
   input Gender $ Internship $ Enrollment $ Count @@;
   datalines;
boys  yes yes 35   boys  yes no 29
boys   no yes 14   boys   no no 27
girls yes yes 32   girls yes no 10
girls  no yes 53   girls  no no 23
;
```

The variable Gender takes the values 'boys' or 'girls,' the variable Internship takes the values 'yes' and 'no,' and the variable Enrollment takes the values 'yes' and 'no.' The variable Count contains the number of students that correspond to each combination of data values. The double at sign (@@) indicates that more than one observation is included on a single data line. In this DATA step, two observations are included on each line.

Researchers are interested in whether there is an association between internship status and summer program enrollment. The Pearson chi-square statistic is an appropriate statistic to assess the association in the corresponding 2×2 table. The following PROC FREQ statements specify this analysis.

You specify the table for which you want to compute statistics with the TABLES statement. You specify the statistics you want to compute with options after a slash (/) in the TABLES statement.

```
proc freq data=SummerSchool order=data;
   tables Internship*Enrollment / chisq;
   weight Count;
run;
```

The ORDER= option controls the order in which variable values are displayed in the rows and columns of the table. By default, the values are arranged according to the alphanumeric order of their unformatted values. If you specify ORDER=DATA, the data are displayed in the same order as they occur in the input data set. Here, because 'yes' appears before 'no' in the data, 'yes' appears first in any table. Other options for controlling order include ORDER=FORMATTED, which orders according to the formatted values, and ORDER=FREQUENCY, which orders by descending frequency count.

In the TABLES statement, Internship*Enrollment specifies a table where the rows are internship status and the columns are program enrollment. The CHISQ option requests chi-square statistics for assessing association between these two variables. Because the input data are in cell count form, the WEIGHT statement is required. The WEIGHT statement names the variable Count, which provides the frequency of each combination of data values.

Figure 36.1 presents the crosstabulation of Internship and Enrollment. In each cell, the values printed under the cell count are the table percentage, row percentage, and column percentage, respectively. For example, in the first cell, 63.21 percent of the students offered courses with internships accepted them and 36.79 percent did not.

Figure 36.1 Crosstabulation Table

```
                    The FREQ Procedure

             Table of Internship by Enrollment

Internship     Enrollment

Frequency|
Percent  |
Row Pct  |
Col Pct  |yes     |no      |  Total
---------+--------+--------+
yes      |     67 |     39 |    106
         |  30.04 |  17.49 |  47.53
         |  63.21 |  36.79 |
         |  50.00 |  43.82 |
---------+--------+--------+
no       |     67 |     50 |    117
         |  30.04 |  22.42 |  52.47
         |  57.26 |  42.74 |
         |  50.00 |  56.18 |
---------+--------+--------+
Total         134       89      223
            60.09    39.91   100.00
```

Figure 36.2 displays the statistics produced by the CHISQ option. The Pearson chi-square statistic is labeled 'Chi-Square' and has a value of 0.8189 with 1 degree of freedom. The associated p-value is 0.3655, which means that there is no significant evidence of an association between internship status and program enrollment. The other chi-square statistics have similar values and are asymptotically equivalent. The other statistics (phi coefficient, contingency coefficient, and Cramer's V) are measures of association derived from the Pearson chi-square. For Fisher's exact test, the two-sided p-value is 0.4122, which also shows no association between internship status and program enrollment.

Figure 36.2 Statistics Produced with the CHISQ Option

Statistic	DF	Value	Prob
Chi-Square	1	0.8189	0.3655
Likelihood Ratio Chi-Square	1	0.8202	0.3651
Continuity Adj. Chi-Square	1	0.5899	0.4425
Mantel-Haenszel Chi-Square	1	0.8153	0.3666
Phi Coefficient		0.0606	
Contingency Coefficient		0.0605	
Cramer's V		0.0606	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	67
Left-sided Pr <= F	0.8513
Right-sided Pr >= F	0.2213
Table Probability (P)	0.0726
Two-sided Pr <= P	0.4122

The analysis, so far, has ignored gender. However, it might be of interest to ask whether program enrollment is associated with internship status after adjusting for gender. You can address this question by doing an analysis of a set of tables (in this case, by analyzing the set consisting of one for boys and one for girls). The Cochran-Mantel-Haenszel (CMH) statistic is appropriate for this situation: it addresses whether rows and columns are associated after controlling for the stratification variable. In this case, you would be stratifying by gender.

The PROC FREQ statements for this analysis are very similar to those for the first analysis, except that there is a third variable, Gender, in the TABLES statement. When you cross more than two variables, the two rightmost variables construct the rows and columns of the table, respectively, and the leftmost variables determine the stratification.

The following PROC FREQ statements also request frequency plots for the crosstabulation tables. PROC FREQ produces these plots by using ODS Graphics to create graphs as part of the procedure output. ODS Graphics must be enabled before producing plots.

```
ods graphics on;
proc freq data=SummerSchool;
   tables Gender*Internship*Enrollment /
          chisq cmh plots(only)=freqplot;
   weight Count;
run;
ods graphics off;
```

This execution of PROC FREQ first produces two individual crosstabulation tables of Internship by Enrollment: one for boys and one for girls. Frequency plots and chi-square statistics are produced for each individual table. Figure 36.3, Figure 36.4, and Figure 36.5 show the results for boys. Note that the chi-square statistic for boys is significant at the $\alpha = 0.05$ level of significance. Boys offered a course with an internship are more likely to enroll than boys who are not.

Figure 36.4 displays the frequency plot of Internship by Enrollment for boys. By default, the frequency plot is displayed as a bar chart with vertical grouping by the row variable Internship. You can use PLOTS= options to request a dot plot instead of a bar chart or to change the orientation of the bars from vertical to horizontal. You can also use PLOTS= options to specify other two-way layouts such as stacked or horizontal grouping.

Figure 36.6, Figure 36.7, and Figure 36.8 display the crosstabulation table, frequency plot, and chi-square statistics for girls. You can see that there is no evidence of association between internship offers and program enrollment for girls.

Figure 36.3 Crosstabulation Table for Boys

The FREQ Procedure

Table 1 of Internship by Enrollment
Controlling for Gender=boys

Internship Enrollment

Frequency Percent Row Pct Col Pct	no	yes	Total
no	27 25.71 65.85 48.21	14 13.33 34.15 28.57	41 39.05
yes	29 27.62 45.31 51.79	35 33.33 54.69 71.43	64 60.95
Total	56 53.33	49 46.67	105 100.00

Figure 36.4 Frequency Plot for Boys

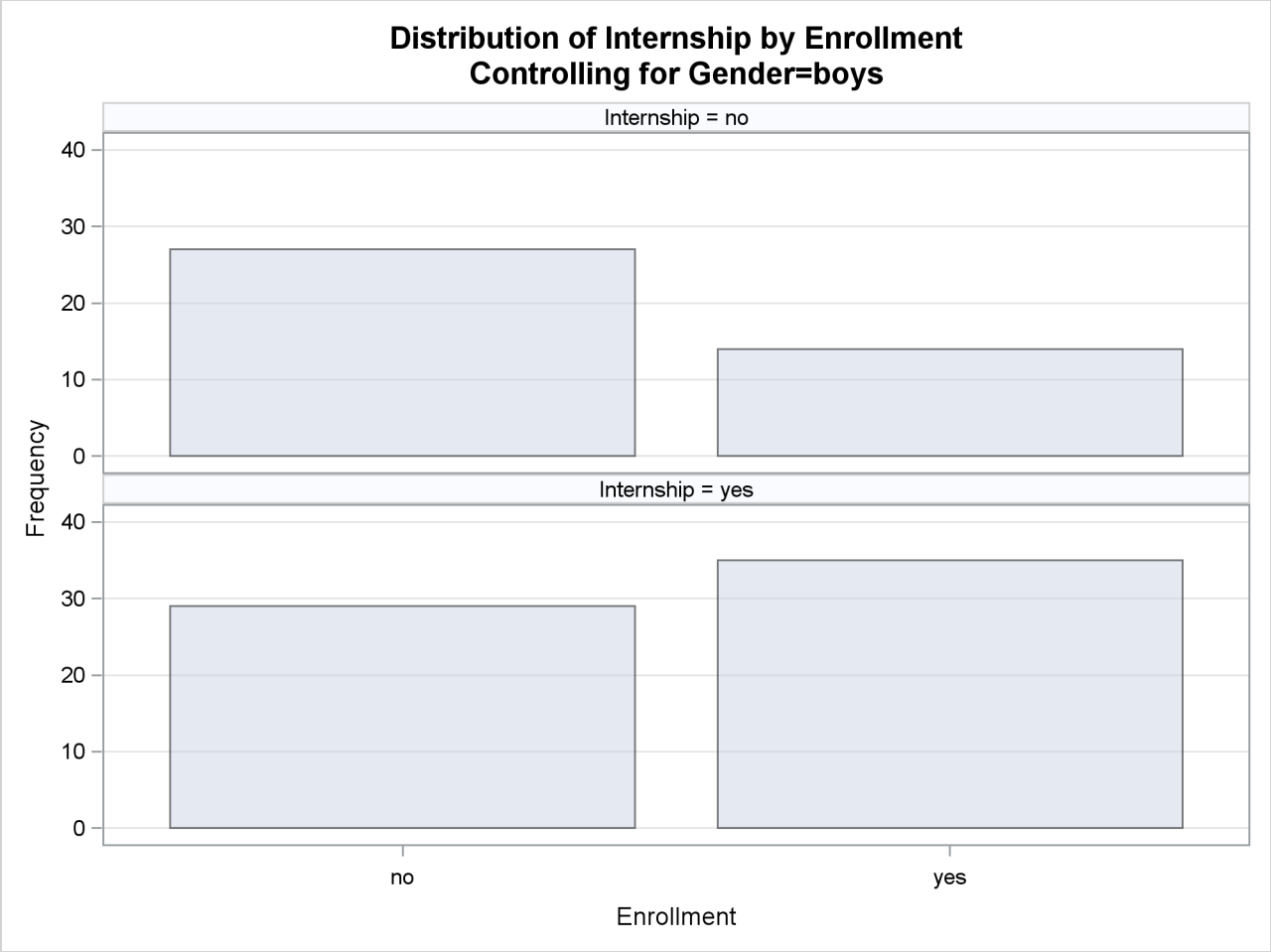


Figure 36.5 Chi-Square Statistics for Boys

Statistic	DF	Value	Prob
Chi-Square	1	4.2366	0.0396
Likelihood Ratio Chi-Square	1	4.2903	0.0383
Continuity Adj. Chi-Square	1	3.4515	0.0632
Mantel-Haenszel Chi-Square	1	4.1963	0.0405
Phi Coefficient		0.2009	
Contingency Coefficient		0.1969	
Cramer's V		0.2009	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	27
Left-sided Pr <= F	0.9885
Right-sided Pr >= F	0.0311
Table Probability (P)	0.0196
Two-sided Pr <= P	0.0467

Figure 36.6 Crosstabulation Table for Girls

```
Table 2 of Internship by Enrollment
    Controlling for Gender=girls

Internship     Enrollment

Frequency|
Percent  |
Row Pct  |
Col Pct  |no      |yes     |  Total
---------+--------+--------+
no       |     23 |     53 |     76
         |  19.49 |  44.92 |  64.41
         |  30.26 |  69.74 |
         |  69.70 |  62.35 |
---------+--------+--------+
yes      |     10 |     32 |     42
         |   8.47 |  27.12 |  35.59
         |  23.81 |  76.19 |
         |  30.30 |  37.65 |
---------+--------+--------+
Total          33       85      118
            27.97    72.03   100.00
```

Figure 36.7 Frequency Plot for Girls

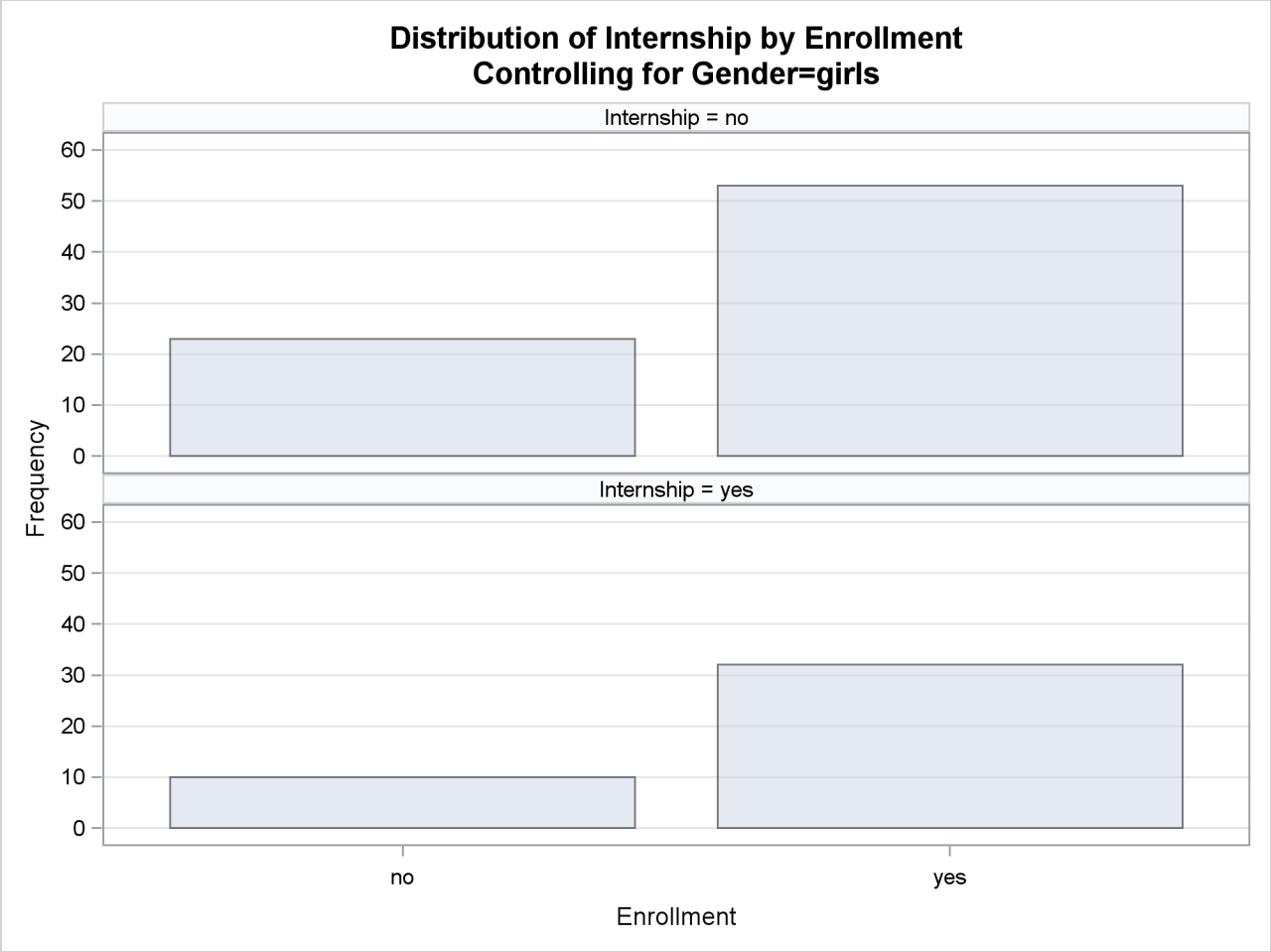


Figure 36.8 Chi-Square Statistics for Girls

Statistic	DF	Value	Prob
Chi-Square	1	0.5593	0.4546
Likelihood Ratio Chi-Square	1	0.5681	0.4510
Continuity Adj. Chi-Square	1	0.2848	0.5936
Mantel-Haenszel Chi-Square	1	0.5545	0.4565
Phi Coefficient		0.0688	
Contingency Coefficient		0.0687	
Cramer's V		0.0688	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	23
Left-sided Pr <= F	0.8317
Right-sided Pr >= F	0.2994
Table Probability (P)	0.1311
Two-sided Pr <= P	0.5245

These individual table results demonstrate the occasional problems with combining information into one table and not accounting for information in other variables such as Gender. Figure 36.9 contains the CMH results. There are three summary (CMH) statistics; which one you use depends on whether your rows and/or columns have an order in $r \times c$ tables. However, in the case of 2×2 tables, ordering does not matter and all three statistics take the same value. The CMH statistic follows the chi-square distribution under the hypothesis of no association, and here, it takes the value 4.0186 with 1 degree of freedom. The associated p-value is 0.0450, which indicates a significant association at the $\alpha = 0.05$ level.

Thus, when you adjust for the effect of gender in these data, there is an association between internship and program enrollment. But, if you ignore gender, no association is found. Note that the CMH option also produces other statistics, including estimates and confidence limits for relative risk and odds ratios for 2×2 tables and the Breslow-Day Test. These results are not displayed here.

Figure 36.9 Test for the Hypothesis of No Association

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	4.0186	0.0450
2	Row Mean Scores Differ	1	4.0186	0.0450
3	General Association	1	4.0186	0.0450

Agreement Study

Medical researchers are interested in evaluating the efficacy of a new treatment for a skin condition. Dermatologists from participating clinics were trained to conduct the study and to evaluate the condition. After the training, two dermatologists examined patients with the skin condition from a pilot study and rated the same patients. The possible evaluations are terrible, poor, marginal, and clear. Table 36.2 contains the data.

Table 36.2 Skin Condition Data

	Dermatologist 2			
Dermatologist 1	**Terrible**	**Poor**	**Marginal**	**Clear**
Terrible	10	4	1	0
Poor	5	10	12	2
Marginal	2	4	12	5
Clear	0	2	6	13

The following DATA step statements create the SAS dataset SkinCondition. The dermatologists' evaluations of the patients are contained in the variables Derm1 and Derm2; the variable Count is the number of patients given a particular pair of ratings.

```
data SkinCondition;
   input Derm1 $ Derm2 $ Count;
   datalines;
terrible terrible 10
terrible     poor 4
terrible marginal 1
terrible    clear 0
poor     terrible 5
poor         poor 10
poor     marginal 12
poor        clear 2
marginal terrible 2
marginal     poor 4
marginal marginal 12
marginal    clear 5
clear    terrible 0
clear        poor 2
clear    marginal 6
clear       clear 13
;
```

The following PROC FREQ statements request an agreement analysis of the skin condition data. In order to evaluate the agreement of the diagnoses (a possible contribution to measurement error in the study), the *kappa coefficient* is computed.

The TABLES statement requests a crosstabulation of the variables Derm1 and Derm2. The AGREE option in the TABLES statement requests the kappa coefficient, together with its standard error and confidence limits. The KAPPA option in the TEST statement requests a test for the null hypothesis that kappa equals zero, or that the agreement is purely by chance. The NOPRINT option in the TABLES statement suppresses the display of the two-way table. The PLOTS= option requests an agreement plot for the two dermatologists. ODS Graphics must be enabled before producing plots.

```
ods graphics on;
proc freq data=SkinCondition order=data;
   tables Derm1*Derm2 /
          agree noprint plots=agreeplot;
   test kappa;
   weight Count;
run;
ods graphics off;
```

Figure 36.10 and Figure 36.11 show the results. The kappa coefficient has the value 0.3449, which indicates some agreement between the dermatologists, and the hypothesis test confirms that you can reject the null hypothesis of no agreement. This conclusion is further supported by the confidence interval of (0.2030, 0.4868), which suggests that the true kappa is greater than zero. The AGREE option also produces Bowker's test for symmetry and the weighted kappa coefficient, but that output is not shown here. Figure 36.11 displays the agreement plot for the ratings of the two dermatologists.

Figure 36.10 Agreement Study

The FREQ Procedure

Statistics for Table of Derm1 by Derm2

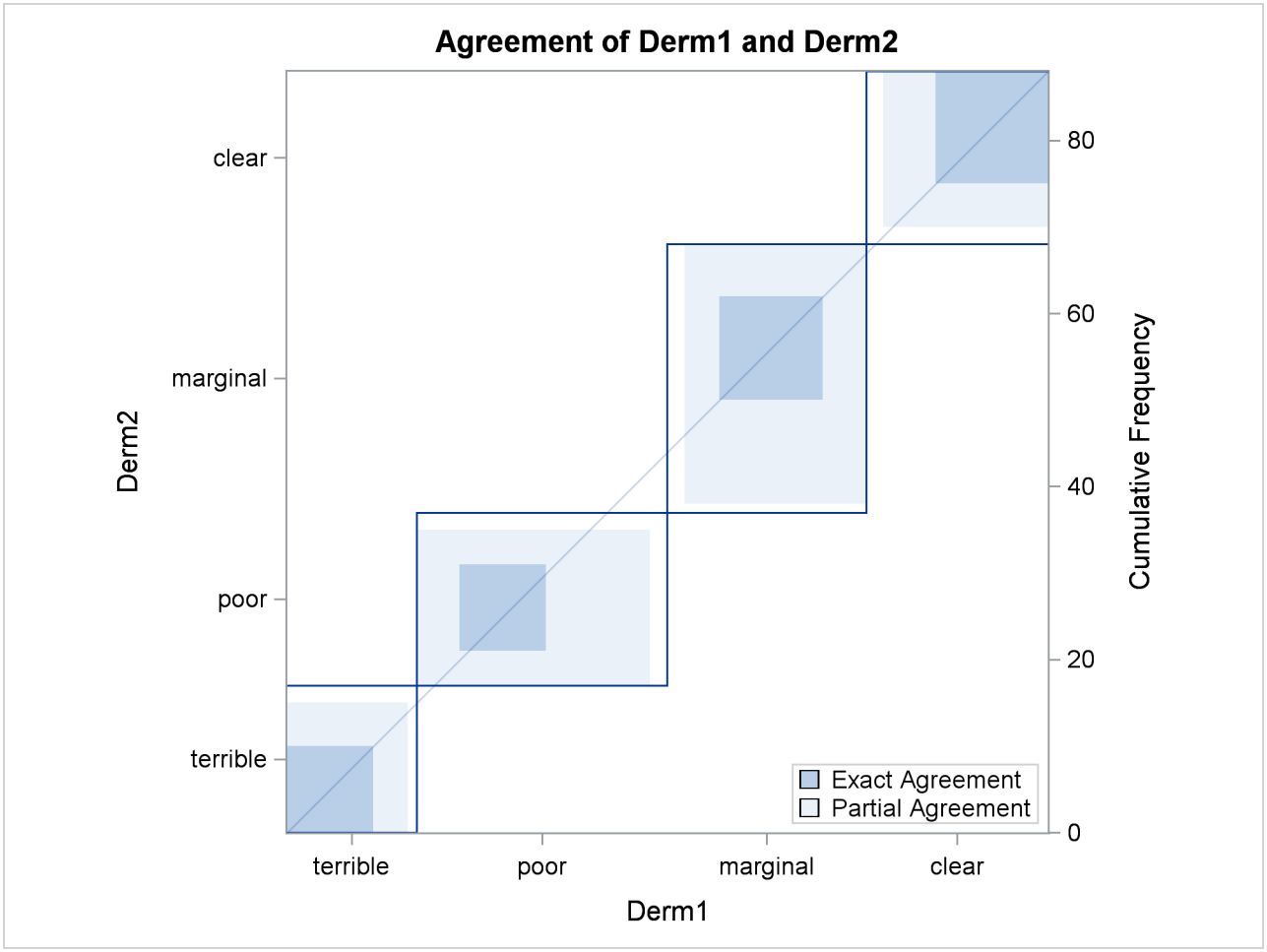
Simple Kappa Coefficient

Kappa	0.3449
ASE	0.0724
95% Lower Conf Limit	0.2030
95% Upper Conf Limit	0.4868

Test of H0: Kappa = 0

ASE under H0	0.0612
Z	5.6366
One-sided Pr > Z	<.0001
Two-sided Pr > |Z|	<.0001

Figure 36.11 Agreement Plot



Syntax: FREQ Procedure

The following statements are available in PROC FREQ:

PROC FREQ < *options* > **;**
BY *variables* **;**
EXACT *statistic-options* < / *computation-options* > **;**
OUTPUT < **OUT=***SAS-data-set* > *options* **;**
TABLES *requests* < / *options* > **;**
TEST *options* **;**
WEIGHT *variable* < / *option* > **;**

The PROC FREQ statement is the only required statement for the FREQ procedure. If you specify the following statements, PROC FREQ produces a one-way frequency table for each variable in the most recently created data set.

```
proc freq;
run;
```

The rest of this section gives detailed syntax information for the BY, EXACT, OUTPUT, TABLES, TEST, and WEIGHT statements in alphabetical order after the description of the PROC FREQ statement. Table 36.3 summarizes the basic function of each PROC FREQ statement.

Table 36.3 Summary of PROC FREQ Statements

Statement	Description
BY	Provides separate analyses for each BY group
EXACT	Requests exact tests
OUTPUT	Requests an output data set
TABLES	Specifies tables and requests analyses
TEST	Requests tests for measures of association and agreement
WEIGHT	Identifies a weight variable

PROC FREQ Statement

PROC FREQ < *options* > **;**

The PROC FREQ statement invokes the procedure and optionally identifies the input data set. By default, the procedure uses the most recently created SAS data set.

Table 36.4 lists the *options* available in the PROC FREQ statement. Descriptions of the *options* follow in alphabetical order.

Table 36.4 PROC FREQ Statement Options

Option	Description
COMPRESS	Begins the next one-way table on the current page
DATA=	Names the input data set
FORMCHAR=	Specifies the outline and cell divider characters for crosstabulation tables
NLEVELS	Displays the number of levels for all TABLES variables
NOPRINT	Suppresses all displayed output
ORDER=	Specifies the order for reporting variable values
PAGE	Displays one table per page

You can specify the following *options* in the PROC FREQ statement.

COMPRESS
begins display of the next one-way frequency table on the same page as the preceding one-way table if there is enough space to begin the table. By default, the next one-way table begins on the current page only if the entire table fits on that page. The COMPRESS option is not valid with the PAGE option.

DATA=*SAS-data-set*
names the SAS data set to be analyzed by PROC FREQ. If you omit the DATA= option, the procedure uses the most recently created SAS data set.

FORMCHAR(1,2,7)=*'formchar-string'*
defines the characters to be used for constructing the outlines and dividers for the cells of crosstabulation table displays. The *formchar-string* should be three characters long. The characters are used to draw the vertical separators (1), the horizontal separators (2), and the vertical-horizontal intersections (7). If you do not specify the FORMCHAR= option, PROC FREQ uses FORMCHAR(1,2,7)='|-+' by default. Table 36.5 summarizes the formatting characters used by PROC FREQ.

Table 36.5 Formatting Characters Used by PROC FREQ

Position	Default	Used to Draw
1	|	Vertical separators
2	-	Horizontal separators
7	+	Intersections of vertical and horizontal separators

The FORMCHAR= option can specify 20 different SAS formatting characters used to display output; however, PROC FREQ uses only the first, second, and seventh formatting characters. Therefore, the proper specification for PROC FREQ is FORMCHAR(1,2,7)= '*formchar-string*'.

Specifying all blanks for *formchar-string* produces crosstabulation tables with no outlines or dividers—for example, FORMCHAR(1,2,7)=' '. You can use any character in *formchar-string*, including hexadecimal characters. If you use hexadecimal characters, you must put an *x* after the closing

quote. For information about which hexadecimal codes to use for which characters, see the documentation for your hardware.

See the CALENDAR, PLOT, and TABULATE procedures in the *Base SAS Procedures Guide* for more information about form characters.

NLEVELS

displays the "Number of Variable Levels" table, which provides the number of levels for each variable named in the TABLES statements. See the section "Number of Variable Levels Table" on page 2374 for details. PROC FREQ determines the variable levels from the formatted variable values, as described in the section "Grouping with Formats" on page 2309.

NOPRINT

suppresses the display of all output. You can use the NOPRINT option when you only want to create an output data set. See the section "Output Data Sets" on page 2371 for information about the output data sets produced by PROC FREQ. Note that the NOPRINT option temporarily disables the Output Delivery System (ODS). For more information, see Chapter 20, "Using the Output Delivery System."

NOTE: A NOPRINT option is also available in the TABLES statement. It suppresses display of the crosstabulation tables but allows display of the requested statistics.

ORDER=DATA | FORMATTED | FREQ | INTERNAL

specifies the order of the variable levels in the frequency and crosstabulation tables, which you request in the TABLES statement.

The ORDER= option can take the following values:

Value of ORDER=	Levels Ordered By
DATA	Order of appearance in the input data set
FORMATTED	External formatted value, except for numeric variables with no explicit format, which are sorted by their unformatted (internal) value
FREQ	Descending frequency count; levels with the most observations come first in the order
INTERNAL	Unformatted value

By default, ORDER=INTERNAL. The FORMATTED and INTERNAL orders are machine-dependent. The ORDER= option does not apply to missing values, which are always ordered first.

For more information about sorting order, see the chapter on the SORT procedure in the *Base SAS Procedures Guide* and the discussion of BY-group processing in *SAS Language Reference: Concepts*.

PAGE

displays only one table per page. Otherwise, PROC FREQ displays multiple tables per page as space permits. The PAGE option is not valid with the COMPRESS option.

BY Statement

BY *variables* **;**

You can specify a BY statement with PROC FREQ to obtain separate analyses on observations in groups defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the BY statement option NOTSORTED or DESCENDING in the BY statement for the FREQ procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure.

For more information about the BY statement, see *SAS Language Reference: Concepts*. For more information about the DATASETS procedure, see the *Base SAS Procedures Guide*.

EXACT Statement

EXACT *statistic-options* < / *computation-options* > **;**

The EXACT statement requests exact tests or confidence limits for the specified statistics. Optionally, PROC FREQ computes Monte Carlo estimates of the exact *p*-values. The *statistic-options* specify the statistics to provide exact tests or confidence limits for. The *computation-options* specify options for the computation of exact statistics. See the section "Exact Statistics" on page 2366 for details.

NOTE: PROC FREQ computes exact tests with fast and efficient algorithms that are superior to direct enumeration. Exact tests are appropriate when a data set is small, sparse, skewed, or heavily tied. For some large problems, computation of exact tests might require a considerable amount of time and memory. Consider using asymptotic tests for such problems. Alternatively, when asymptotic methods might not be sufficient for such large problems, consider using Monte Carlo estimation of exact *p*-values. You can request Monte Carlo estimation by specifying the MC *computation-option* in the EXACT statement. See the section "Computational Resources" on page 2369 for more information.

Statistic Options

The *statistic-options* specify the statistics to provide exact tests or confidence limits for.

For one-way tables, exact *p*-values are available for the binomial proportion tests and the chi-square goodness-of-fit test. Exact (Clopper-Pearson) confidence limits are available for the binomial proportion.

For two-way tables, exact p-values are available for the following tests: Pearson chi-square test, likelihood-ratio chi-square test, Mantel-Haenszel chi-square test, Fisher's exact test, Jonckheere-Terpstra test, and Cochran-Armitage test for trend. Exact p-values are also available for tests of the following statistics: Pearson correlation coefficient, Spearman correlation coefficient, Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$, Somers' $D(R|C)$, simple kappa coefficient, and weighted kappa coefficient.

For 2×2 tables, PROC FREQ provides McNemar's exact test and exact confidence limits for the odds ratio. PROC FREQ also provides exact unconditional confidence limits for the risk (proportion) difference and for the relative risk (ratio of proportions). For stratified 2×2 tables, PROC FREQ provides Zelen's exact test for equal odds ratios, exact confidence limits for the common odds ratio, and an exact test for the common odds ratio.

Table 36.6 lists the available *statistic-options* and the exact statistics computed. For more information about these statistics, see the TABLES statement and the section "Statistical Computations" on page 2314. For more information about exact computations, see the section "Exact Statistics" on page 2366.

Most of the option names listed in Table 36.6 are identical to the corresponding option names in the TABLES and OUTPUT statements. You can request exact computations for groups of statistics by using options that are identical to the following TABLES statement options: CHISQ, MEASURES, and AGREE. For example, when you specify the CHISQ option in the EXACT statement, PROC FREQ computes exact p-values for the Pearson chi-square, likelihood-ratio chi-square, and Mantel-Haenszel chi-square tests. You can request exact computations for an individual statistic by specifying the corresponding *statistic-option* from the list in Table 36.6.

Table 36.6 EXACT Statement Statistic Options

Statistic Option	Exact Statistics
AGREE	McNemar's test (for 2×2 tables), simple kappa test, weighted kappa test
BINOMIAL	Binomial proportion tests for one-way tables
CHISQ	Chi-square goodness-of-fit test for one-way tables; Pearson chi-square, likelihood-ratio chi-square, and Mantel-Haenszel chi-square tests for two-way tables
COMOR	Confidence limits for the common odds ratio, common odds ratio test (for $h \times 2 \times 2$ tables)
EQOR | ZELEN	Zelen's test for equal odds ratios (for $h \times 2 \times 2$ tables)
FISHER	Fisher's exact test
JT	Jonckheere-Terpstra test
KAPPA	Test for the simple kappa coefficient
KENTB	Test for Kendall's tau-b
LRCHI	Likelihood-ratio chi-square test
MCNEM	McNemar's test (for 2×2 tables)
MEASURES	Tests for the Pearson correlation and Spearman correlation, confidence limits for the odds ratio (for 2×2 tables)
MHCHI	Mantel-Haenszel chi-square test
OR	Confidence limits for the odds ratio (for 2×2 tables)
PCHI	Pearson chi-square test
PCORR	Test for the Pearson correlation coefficient

Table 36.6 *continued*

Statistic Option	Exact Statistics
RELRISK	Confidence limits for the relative risk (for 2 × 2 tables)
RISKDIFF	Confidence limits for the proportion difference (for 2×2 tables)
SCORR	Test for the Spearman correlation coefficient
SMDCR	Test for Somers' $D(C\|R)$
SMDRC	Test for Somers' $D(R\|C)$
STUTC	Test for Stuart's tau-c
TREND	Cochran-Armitage test for trend
WTKAP	Test for the weighted kappa coefficient

You can specify *options* for the following two EXACT statement *statistic-options*:

RELRISK < (*options*) >
requests exact unconditional confidence limits for the relative risk for 2 × 2 tables. PROC FREQ computes the confidence limits by inverting two separate one-sided exact tests (Santner and Snell 1980). By default, this computation uses the unstandardized relative risk as the test statistic. If you specify the RELRISK(METHOD=FMSCORE) option, PROC FREQ uses the Farrington-Manning score statistic (Chan and Zhang 1999). See the section "Exact Unconditional Confidence Limits for the Relative Risk" on page 2348 for more information.

You can set the confidence level by using the ALPHA= option in the TABLES statement. The default of ALPHA=0.5 produces 95% confidence limits.

You can specify the following *options* inside parentheses after the RELRISK *statistic-option*:

COLUMN=1 | 2 | BOTH
specifies the 2 × 2 table column for which to compute the relative risk. The default is COLUMN=1, which provides exact confidence limits for the column 1 relative risk. If you specify COLUMN=BOTH, PROC FREQ provides exact confidence limits for both column 1 and column 2 relative risks.

METHOD=FMSCORE | SCORE
requests exact unconditional confidence limits that are based on the Farrington-Manning score statistic (Chan and Zhang 1999). See the section "Exact Unconditional Confidence Limits for the Relative Risk" on page 2348 for more information. If you do not specify METHOD=FMSCORE, by default PROC FREQ uses the unstandardized relative risk in the exact confidence limit computations.

RISKDIFF < (*options*) >
requests exact unconditional confidence limits for the risk difference for 2 × 2 tables. PROC FREQ computes the confidence limits by inverting two separate one-sided exact tests (Santner and Snell 1980). By default, this computation uses the unstandardized risk difference as the test statistic. If you specify the RISKDIFF(METHOD=FMSCORE) option, PROC FREQ uses the Farrington-Manning score statistic (Chan and Zhang 1999). See the section "Exact Unconditional Confidence Limits for the Risk Difference" on page 2345 for more information.

You can set the confidence level by using the ALPHA= option in the TABLES statement. The default of ALPHA=0.5 produces 95% confidence limits.

You can specify the following *options* inside parentheses after the RISKDIFF *statistic-option*:

COLUMN=1 | 2 | BOTH
specifies the 2×2 table column for which to compute the risk difference. The default is COLUMN=BOTH, which provides exact confidence limits for both column 1 and column 2 risk differences.

METHOD=FMSCORE | SCORE
requests exact unconditional confidence limits that are based on the Farrington-Manning score statistic (Chan and Zhang 1999). See the section "Exact Unconditional Confidence Limits for the Risk Difference" on page 2345 for more information. If you do not specify METHOD=FMSCORE, by default PROC FREQ uses the unstandardized risk difference in the exact confidence limit computations.

Using TABLES Statement Options with the EXACT Statement

If you use only one TABLES statement, you do not need to specify the same options in both the TABLES and EXACT statements; when you specify a *statistic-option* in the EXACT statement, PROC FREQ automatically invokes the corresponding TABLES statement option. However, when you use multiple TABLES statements and want exact computations, you must specify options in the TABLES statements to request the desired statistics. PROC FREQ then performs exact computations for all statistics that you also specify in the EXACT statement.

The TABLES statement group option CHISQ includes tests that correspond to the following EXACT statement individual *statistic-options*: LRCHI, MHCHI, and PCHI. The MEASURES option in the TABLES statement includes statistics that correspond to the following EXACT statement *statistic-options*: KENTB, OR, PCORR, SCORR, SMDCR, SMDRC, and STUTC. The AGREE option in the TABLES statement produces analyses that correspond to the KAPPA, MCNEM, and WTKAP *statistic-options* in the EXACT statement. The CMH option in the TABLES statement produces analyses that correspond to the COMOR and EQOR (ZELEN) *statistic-options* in the EXACT statement.

Computation Options

The *computation-options* specify options for computation of exact statistics. You can specify the following *computation-options* in the EXACT statement after a slash (/).

ALPHA=α
specifies the level of the confidence limits for Monte Carlo p-value estimates. The value of α must be between 0 and 1, and the default is 0.01. A confidence level of α produces $100(1-\alpha)\%$ confidence limits. The default of ALPHA=.01 produces 99% confidence limits for the Monte Carlo estimates.

The ALPHA= option invokes the MC option.

MAXTIME=*value*
specifies the maximum clock time (in seconds) that PROC FREQ can use to compute an exact p-value. If the procedure does not complete the computation within the specified time, the computation

terminates. The value of MAXTIME= must be a positive number. The MAXTIME= option is valid for Monte Carlo estimation of exact *p*-values, as well as for direct exact *p*-value computation. See the section "Computational Resources" on page 2369 for more information.

MC

requests Monte Carlo estimation of exact *p*-values instead of direct exact *p*-value computation. Monte Carlo estimation can be useful for large problems that require a considerable amount of time and memory for exact computations but for which asymptotic approximations might not be sufficient. See the section "Monte Carlo Estimation" on page 2369 for more information.

The MC option is available for all EXACT *statistic-options* except the BINOMIAL option and the following options that apply only to 2×2 or $h \times 2 \times 2$ tables: COMOR, EQOR, MCNEM, OR, RELRISK, and RISKDIFF. PROC FREQ computes only exact tests or confidence limits for these statistics.

The ALPHA=, N=, and SEED= options also invoke the MC option.

N=*n*

specifies the number of samples for Monte Carlo estimation. The value of *n* must be a positive integer, and the default is 10,000. Larger values of *n* produce more precise estimates of exact *p*-values. Because larger values of *n* generate more samples, the computation time increases.

The N= option invokes the MC option.

POINT

requests exact point probabilities for the test statistics.

The POINT option is available for all the EXACT statement *statistic-options* except the OR. RELRISK, and RISKDIFF options, which provide exact confidence limits. The POINT option is not available with the MC option.

SEED=*number*

specifies the initial seed for random number generation for Monte Carlo estimation. The value of the SEED= option must be an integer. If you do not specify the SEED= option or if the SEED= value is negative or zero, PROC FREQ uses the time of day from the computer's clock to obtain the initial seed.

The SEED= option invokes the MC option.

OUTPUT Statement

OUTPUT < *OUT= SAS-data-set* > *options* **;**

The OUTPUT statement creates a SAS data set that contains statistics computed by PROC FREQ. You specify which statistics to store in the output data set with the OUTPUT statement *options*. The output data set contains one observation for each two-way table or stratum, and one observation for summary statistics across all strata. For more information about the contents of the output data set, see the section "Contents of the OUTPUT Statement Output Data Set" on page 2373.

Only one OUTPUT statement is allowed for each execution of PROC FREQ. You must specify a TABLES statement with the OUTPUT statement. If you use multiple TABLES statements, the contents of the OUTPUT data set correspond to the last TABLES statement. If you use multiple table requests in a TABLES statement, the contents of the OUTPUT data set correspond to the last table request.

Note that you can use the Output Delivery System (ODS) to create a SAS data set from any piece of PROC FREQ output. For more information, see the section "ODS Table Names" on page 2382.

Also note that the output data set created by the OUTPUT statement is not the same as the output data set created by the OUT= option in the TABLES statement. The OUTPUT statement creates a data set that contains statistics (such as the Pearson chi-square and its p-value), and the OUT= option in the TABLES statement creates a data set that contains frequency table counts and percentages. See the section "Output Data Sets" on page 2371 for more information.

You can specify the following *options* in the OUTPUT statement:

OUT=*SAS-data-set*
names the output data set. If you omit the OUT= option, the data set is named DATAn, where n is the smallest integer that makes the name unique.

options
specify the statistics you want in the output data set. Table 36.7 lists the available *options*, together with the TABLES statement options needed to produce the statistics. You can output groups of statistics by using group options identical to those available in the TABLES statement, which include the AGREE, ALL, CHISQ, CMH, and MEASURES options. Or you can request statistics individually.

When you specify an option in the OUTPUT statement, the output data set contains all statistics from that analysis—the estimate or test statistic plus any associated standard error, confidence limits, p-values, and degrees of freedom. See the section "Contents of the OUTPUT Statement Output Data Set" on page 2373 for details.

If you want to store a statistic in the output data set, you must also request computation of that statistic with the appropriate TABLES or EXACT statement option. For example, you cannot specify the option PCHI (Pearson chi-square) in the OUTPUT statement without also specifying a TABLES or EXACT statement option to compute the Pearson chi-square test. The TABLES statement option ALL or CHISQ requests the Pearson chi-square test. If you have only one TABLES statement, the EXACT statement option CHISQ or PCHI also requests the Pearson chi-square test. Table 36.7 lists the TABLES statement *options* required to produce the OUTPUT data set statistics. Note that the ALL option in the TABLES statement invokes the CHISQ, MEASURES, and CMH options.

Table 36.7 OUTPUT Statement Options

Option	Output Data Set Statistics	Required TABLES Statement Option
AGREE	McNemar's test, Bowker's test, simple and weighted kappas; for multiple strata, overall simple and weighted kappas, tests for equal kappas, and Cochran's Q	AGREE
AJCHI	Continuity-adjusted chi-square (2×2 tables)	CHISQ
ALL	CHISQ, MEASURES, and CMH statistics; N	ALL

Table 36.7 *continued*

Option	Output Data Set Statistics	Required TABLES Statement Option
BDCHI	Breslow-Day test ($h \times 2 \times 2$ tables)	CMH, CMH1, or CMH2
BINOMIAL	Binomial statistics for one-way tables	BINOMIAL
CHISQ	For one-way tables, goodness-of-fit test; for two-way tables, Pearson, likelihood-ratio, continuity-adjusted, and Mantel-Haenszel chi-squares, Fisher's exact test (2×2 tables), phi and contingency coefficients, Cramer's V	CHISQ
CMH	Cochran-Mantel-Haenszel (CMH) correlation, row mean scores (ANOVA), and general association statistics; for 2×2 tables, logit and Mantel-Haenszel adjusted odds ratios and relative risks, Breslow-Day test	CMH
CMH1	CMH output, except row mean scores (ANOVA) and general association statistics	CMH or CMH1
CMH2	CMH output, except general association statistic	CMH or CMH2
CMHCOR	CMH correlation statistic	CMH, CMH1, or CMH2
CMHGA	CMH general association statistic	CMH
CMHRMS	CMH row mean scores (ANOVA) statistic	CMH or CMH2
COCHQ	Cochran's Q ($h \times 2 \times 2$ tables)	AGREE
CONTGY	Contingency coefficient	CHISQ
CRAMV	Cramer's V	CHISQ
EQKAP	Test for equal simple kappas	AGREE
EQOR | ZELEN	Zelen's test for equal odds ratios ($h \times 2 \times 2$ tables)	CMH and EXACT EQOR
EQWKP	Test for equal weighted kappas	AGREE
FISHER	Fisher's exact test	CHISQ or FISHER [1]
GAILSIMON	Gail-Simon test	GAILSIMON
GAMMA	Gamma	MEASURES
JT	Jonckheere-Terpstra test	JT
KAPPA	Simple kappa coefficient	AGREE
KENTB	Kendall's tau-b	MEASURES
LAMCR	Lambda asymmetric ($C\,|\,R$)	MEASURES
LAMDAS	Lambda symmetric	MEASURES
LAMRC	Lambda asymmetric ($R|C$)	MEASURES
LGOR	Adjusted logit odds ratio ($h \times 2 \times 2$ tables)	CMH, CMH1, or CMH2
LGRRC1	Adjusted column 1 logit relative risk	CMH, CMH1, or CMH2
LGRRC2	Adjusted column 2 logit relative risk	CMH, CMH1, or CMH2
LRCHI	Likelihood-ratio chi-square	CHISQ
MCNEM	McNemar's test (2×2 tables)	AGREE

[1] CHISQ computes Fisher's exact test for 2×2 tables. Use the FISHER option to compute Fisher's exact test for general $r \times c$ tables.

Table 36.7 *continued*

Option	Output Data Set Statistics	Required TABLES Statement Option
MEASURES	Gamma, Kendall's tau-b, Stuart's tau-c, Somers' $D(C\|R)$ and $D(R\|C)$, Pearson and Spearman correlations, lambda asymmetric $(C\|R)$ and $(R\|C)$, lambda symmetric, uncertainty coefficients $(C\|R)$ and $(R\|C)$, symmetric uncertainty coefficient; odds ratio and relative risks (2×2 tables)	MEASURES
MHCHI	Mantel-Haenszel chi-square	CHISQ
MHOR | COMOR	Adjusted Mantel-Haenszel odds ratio ($h \times 2 \times 2$ tables)	CMH, CMH1, or CMH2
MHRRC1	Adjusted column 1 Mantel-Haenszel relative risk	CMH, CMH1, or CMH2
MHRRC2	Adjusted column 2 Mantel-Haenszel relative risk	CMH, CMH1, or CMH2
N	Number of nonmissing observations	
NMISS	Number of missing observations	
OR	Odds ratio (2×2 tables)	MEASURES or RELRISK
PCHI	Chi-square goodness-of-fit test for one-way tables, Pearson chi-square for two-way tables	CHISQ
PCORR	Pearson correlation coefficient	MEASURES
PHI	Phi coefficient	CHISQ
PLCORR	Polychoric correlation coefficient	PLCORR
RDIF1	Column 1 risk difference (row 1 - row 2)	RISKDIFF
RDIF2	Column 2 risk difference (row 1 - row 2)	RISKDIFF
RELRISK	Odds ratio and relative risks (2×2 tables)	MEASURES or RELRISK
RISKDIFF	Risks and risk differences (2×2 tables)	RISKDIFF
RISKDIFF1	Column 1 risks and risk difference	RISKDIFF
RISKDIFF2	Column 2 risks and risk difference	RISKDIFF
RRC1	Column 1 relative risk	MEASURES or RELRISK
RRC2	Column 2 relative risk	MEASURES or RELRISK
RSK1	Column 1 risk, overall	RISKDIFF
RSK11	Column 1 risk, for row 1	RISKDIFF
RSK12	Column 2 risk, for row 1	RISKDIFF
RSK2	Column 2 risk, overall	RISKDIFF
RSK21	Column 1 risk, for row 2	RISKDIFF
RSK22	Column 2 risk, for row 2	RISKDIFF
SCORR	Spearman correlation coefficient	MEASURES
SMDCR	Somers' $D(C\|R)$	MEASURES
SMDRC	Somers' $D(R\|C)$	MEASURES
STUTC	Stuart's tau-c	MEASURES
TREND	Cochran-Armitage test for trend	TREND
TSYMM	Bowker's test of symmetry	AGREE
U	Symmetric uncertainty coefficient	MEASURES
UCR	Uncertainty coefficient $(C\|R)$	MEASURES
URC	Uncertainty coefficient $(R\|C)$	MEASURES
WTKAP	Weighted kappa coefficient	AGREE

TABLES Statement

TABLES *requests* < / *options* > **;**

The TABLES statement requests one-way to *n*-way frequency and crosstabulation tables and statistics for those tables.

If you omit the TABLES statement, PROC FREQ generates one-way frequency tables for all data set variables that are not listed in the other statements.

The following argument is required in the TABLES statement.

requests

specify the frequency and crosstabulation tables to produce. A request is composed of one variable name or several variable names separated by asterisks. To request a one-way frequency table, use a single variable. To request a two-way crosstabulation table, use an asterisk between two variables. To request a multiway table (an *n*-way table, where *n*>2), separate the desired variables with asterisks. The unique values of these variables form the rows, columns, and strata of the table. You can include up to 50 variables in a single multiway table request.

For two-way to multiway tables, the values of the last variable form the crosstabulation table columns, while the values of the next-to-last variable form the rows. Each level (or combination of levels) of the other variables forms one stratum. PROC FREQ produces a separate crosstabulation table for each stratum. For example, a specification of A*B*C*D in a TABLES statement produces *k* tables, where *k* is the number of different combinations of values for A and B. Each table lists the values for C down the side and the values for D across the top.

You can use multiple TABLES statements in the PROC FREQ step. PROC FREQ builds all the table requests in one pass of the data, so that there is essentially no loss of efficiency. You can also specify any number of table requests in a single TABLES statement. To specify multiple table requests quickly, use a grouping syntax by placing parentheses around several variables and joining other variables or variable combinations. For example, the statements shown in Table 36.8 illustrate grouping syntax.

Table 36.8 Grouping Syntax

TABLES Request	Equivalent to
A*(B C)	A*B A*C
(A B)*(C D)	A*C B*C A*D B*D
(A B C)*D	A*D B*D C*D
A – – C	A B C
(A – – C)*D	A*D B*D C*D

The TABLES statement variables are one or more variables from the DATA= input data set. These variables can be either character or numeric, but the procedure treats them as categorical variables. PROC FREQ uses the formatted values of the TABLES variable to determine the categorical variable levels. So if you assign a format to a variable with a FORMAT statement, PROC FREQ formats the values before dividing observations into the levels of a frequency or crosstabulation table. See the discussion of the FORMAT procedure in the *Base SAS Procedures Guide* and the discussions of the FORMAT statement and SAS formats in *SAS Language Reference: Dictionary*.

If you use PROC FORMAT to create a user-written format that combines missing and nonmissing values into one category, PROC FREQ treats the entire category of formatted values as missing. See the discussion in the section "Grouping with Formats" on page 2309 for more information.

By default, the frequency or crosstabulation table lists the values of both character and numeric variables in ascending order based on internal (unformatted) variable values. You can change the order of the values in the table by specifying the ORDER= option in the PROC FREQ statement. To list the values in ascending order by formatted value, use ORDER=FORMATTED.

Without Options

If you request a one-way frequency table for a variable without specifying options, PROC FREQ produces frequencies, cumulative frequencies, percentages of the total frequency, and cumulative percentages for each value of the variable. If you request a two-way or an *n*-way crosstabulation table without specifying options, PROC FREQ produces crosstabulation tables that include cell frequencies, cell percentages of the total frequency, cell percentages of row frequencies, and cell percentages of column frequencies. The procedure excludes observations with missing values from the table but displays the total frequency of missing observations below each table.

Options

Table 36.9 lists the *options* available in the TABLES statement. Descriptions of the *options* follow in alphabetical order.

Table 36.9 TABLES Statement Options

Option	Description
Control Statistical Analysis	
AGREE	Requests tests and measures of classification agreement
ALL	Requests tests and measures of association produced by CHISQ, MEASURES, and CMH
ALPHA=	Sets the confidence level for confidence limits
BDT	Requests Tarone's adjustment for the Breslow-Day test
BINOMIAL	Requests binomial proportion, confidence limits, and tests for one-way tables
BINOMIALC	Requests BINOMIAL statistics with a continuity correction
CHISQ	Requests chi-square tests and measures based on chi-square
CL	Requests confidence limits for the MEASURES statistics
CMH	Requests all Cochran-Mantel-Haenszel statistics
CMH1	Requests CMH correlation statistic, adjusted odds ratios, and adjusted relative risks
CMH2	Requests CMH correlation and row mean scores (ANOVA) statistics, adjusted odds ratios, and adjusted relative risks
CONVERGE=	Specifies convergence criterion for polychoric correlation
FISHER	Requests Fisher's exact test for tables larger than 2×2
GAILSIMON	Requests Gail-Simon test for qualitative interactions
JT	Requests Jonckheere-Terpstra test
MAXITER=	Specifies maximum number of iterations for polychoric correlation
MEASURES	Requests measures of association

Table 36.9 *continued*

Option	Description
MISSING	Treats missing values as nonmissing
PLCORR	Requests polychoric correlation
RELRISK	Requests relative risk measures for 2×2 tables
RISKDIFF	Requests risks and risk differences for 2×2 tables
SCORES=	Specifies the type of row and column scores
TESTF=	Specifies expected frequencies for one-way chi-square test
TESTP=	Specifies expected proportions for one-way chi-square test
TREND	Requests Cochran-Armitage test for trend
Control Additional Table Information	
CELLCHI2	Displays cell contributions to the Pearson chi-square statistic
CUMCOL	Displays cumulative column percentages
DEVIATION	Displays deviations of cell frequencies from expected values
EXPECTED	Displays expected cell frequencies
MISSPRINT	Displays missing value frequencies
SPARSE	Includes all possible combinations of variable levels in LIST and OUT=
TOTPCT	Displays percentages of total frequency for n-way tables ($n > 2$)
Control Displayed Output	
CONTENTS=	Specifies the contents label for crosstabulation tables
CROSSLIST	Displays crosstabulation tables in ODS column format
FORMAT=	Formats the frequencies in crosstabulation tables
LIST	Displays two-way to n-way tables in list format
NOCOL	Suppresses display of column percentages
NOCUM	Suppresses display of cumulative frequencies and percentages
NOFREQ	Suppresses display of frequencies
NOPERCENT	Suppresses display of percentages
NOPRINT	Suppresses display of crosstabulation tables but displays statistics
NOROW	Suppresses display of row percentages
NOSPARSE	Suppresses zero frequency levels in CROSSLIST, LIST and OUT=
NOWARN	Suppresses log warning message for the chi-square test
PRINTKWT	Displays kappa coefficient weights
SCOROUT	Displays row and column scores
Produce Statistical Graphics	
PLOTS=	Requests plots from ODS Graphics
Create an Output Data Set	
OUT=	Names an output data set to contain frequency counts
OUTCUM	Includes cumulative frequencies and percentages in the output data set for one-way tables
OUTEXPECT	Includes expected frequencies in the output data set
OUTPCT	Includes row, column, and two-way table percentages in the output data set

You can specify the following *options* in a TABLES statement.

AGREE < (WT=FC) >

requests tests and measures of classification agreement for square tables. The AGREE option provides McNemar's test for 2×2 tables and Bowker's test of symmetry for square tables with more than two response categories. The AGREE option also produces the simple kappa coefficient, the weighted kappa coefficient, their asymptotic standard errors, and their confidence limits. When there are multiple strata, the AGREE option provides overall simple and weighted kappas as well as tests for equal kappas among strata. When there are multiple strata and two response categories, PROC FREQ computes Cochran's Q test. See the section "Tests and Measures of Agreement" on page 2352 for details about these statistics.

If you specify the WT=FC option in parentheses following the AGREE option, PROC FREQ uses Fleiss-Cohen weights to compute the weighted kappa coefficient. By default, PROC FREQ uses Cicchetti-Allison weights. See the section "Weighted Kappa Coefficient" on page 2354 for details. You can specify the PRINTKWT option to display the kappa coefficient weights.

AGREE statistics are computed only for square tables, where the number of rows equals the number of columns. If your table is not square due to observations with zero weights, you can specify the ZEROS option in the WEIGHT statement to include these observations. For more details, see the section "Tables with Zero Rows and Columns" on page 2356.

You can use the TEST statement to request asymptotic tests for the simple and weighted kappa coefficients. You can request exact p-values for the simple and weighted kappa coefficient tests, as well as for McNemar's test, by specifying the corresponding options in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

ALL

requests all of the tests and measures that are computed by the CHISQ, MEASURES, and CMH options. The number of CMH statistics computed can be controlled by the CMH1 and CMH2 options.

ALPHA=α

specifies the level of confidence limits. The value of α must be between 0 and 1, and the default is 0.05. A confidence level of α produces $100(1-\alpha)\%$ confidence limits. The default of ALPHA=0.05 produces 95% confidence limits.

ALPHA= applies to confidence limits requested by TABLES statement options. There is a separate ALPHA= option in the EXACT statement that sets the level of confidence limits for Monte Carlo estimates of exact p-values, which are requested in the EXACT statement.

BDT

requests Tarone's adjustment in the Breslow-Day test for homogeneity of odds ratios. (You must specify the CMH option to compute the Breslow-Day test.) See the section "Breslow-Day Test for Homogeneity of the Odds Ratios" on page 2363 for more information.

BINOMIAL < (*binomial-options*) >

requests the binomial proportion for one-way tables. When you specify the BINOMIAL option, by default PROC FREQ also provides the asymptotic standard error, asymptotic (Wald) and exact (Clopper-Pearson) confidence limits, and the asymptotic equality test for the binomial proportion.

You can specify *binomial-options* inside parentheses following the BINOMIAL option. The LEVEL= *binomial-option* identifies the variable level for which to compute the proportion. If you do not specify LEVEL=, PROC FREQ computes the proportion for the first level that appears in the output. The P= *binomial-option* specifies the null proportion for the binomial tests. If you do not specify P=, PROC FREQ uses P=0.5 by default.

You can also specify *binomial-options* to request additional tests and confidence limits for the binomial proportion. The EQUIV, NONINF, and SUP *binomial-options* request tests of equivalence, noninferiority, and superiority, respectively. Table 36.10 summarizes the *binomial-options*.

Available confidence limits for the binomial proportion include Agresti-Coull, exact (Clopper-Pearson), Jeffreys, Wald, and Wilson (score) confidence limits. You can specify more than one type of binomial confidence limits in the same analysis. If you do not specify any confidence limit requests with *binomial-options*, PROC FREQ computes Wald asymptotic confidence limits and exact (Clopper-Pearson) confidence limits by default. The ALPHA= option determines the confidence level, and the default of ALPHA=0.05 produces 95% confidence limits for the binomial proportion.

As part of the noninferiority, superiority, and equivalence analyses, PROC FREQ provides test-based confidence limits that have a confidence coefficient of $100(1 - 2\alpha)\%$ (Schuirmann 1999). The ALPHA= option determines the confidence level, and the default of ALPHA=0.05 produces 90% confidence limits. See the sections "Noninferiority Test" on page 2333 and "Equivalence Test" on page 2335 for details.

To request exact tests for the binomial proportion, specify the BINOMIAL option in the EXACT statement. PROC FREQ then computes exact p-values for all binomial tests that you request with *binomial-options*, which can include tests of noninferiority, superiority, and equivalence, in addition to the test of equality.

See the section "Binomial Proportion" on page 2329 for details.

Table 36.10 BINOMIAL Options

Option	Description
LEVEL=	Specifies the variable level
P=	Specifies the null proportion
CORRECT	Requests continuity correction
Request Confidence Limits	
AGRESTICOULL | AC	Requests Agresti-Coull confidence limits
ALL	Requests all confidence limits
EXACT | CLOPPERPEARSON	Requests Clopper-Pearson confidence limits
JEFFREYS | J	Requests Jeffreys confidence limits
WALD	Requests Wald confidence limits
WILSON | W	Requests Wilson (score) confidence limits
Request Tests	
EQUIV | EQUIVALENCE	Requests an equivalence test
NONINF | NONINFERIORITY	Requests a noninferiority test
SUP | SUPERIORITY	Requests a superiority test
MARGIN=	Specifies the test margin
VAR=SAMPLE | NULL	Specifies the test variance

You can specify the following *binomial-options* inside parentheses following the BINOMIAL option:

AGRESTICOULL | AC
requests Agresti-Coull confidence limits for the binomial proportion. See the section "Agresti-Coull Confidence Limits" on page 2330 for details.

ALL
requests all available types of confidence limits for the binomial proportion. These include the following: Agresti-Coull, exact (Clopper-Pearson), Jeffreys, Wald, and Wilson (score) confidence limits.

CORRECT
includes a continuity correction in the Wald confidence limits and tests. The BINOMIAL(CORRECT) option is equivalent to the BINOMIALC option.

EQUIV | EQUIVALENCE
requests a test of equivalence for the binomial proportion. See the section "Equivalence Test" on page 2335 for details. You can specify the equivalence test margins, the null proportion, and the variance type with the MARGIN=, P=, and VAR= *binomial-options*, respectively.

EXACT | CLOPPERPEARSON
requests exact (Clopper-Pearson) confidence limits for the binomial proportion. See the section "Exact (Clopper-Pearson) Confidence Limits" on page 2331 for details. If you do not request any binomial confidence limits by specifying *binomial-options*, PROC FREQ produces Wald and exact (Clopper-Pearson) confidence limits by default. To request exact tests for the binomial proportion, specify the BINOMIAL option in the EXACT statement.

JEFFREYS | J
requests Jeffreys confidence limits for the binomial proportion. See the section "Jeffreys Confidence Limits" on page 2330 for details.

LEVEL=*level-number* | '*level-value*'
specifies the variable level for the binomial proportion. By default, PROC FREQ computes the proportion of observations for the first variable level that appears in the output. To request a different level, use LEVEL=*level-number* or LEVEL='*level-value*', where *level-number* is the variable level's number or order in the output, and *level-value* is the formatted value of the variable level. The value of *level-number* must be a positive integer. You must enclose *level-value* in single quotes.

MARGIN=*value* | **(***lower***,***upper***)**
specifies the margin for the noninferiority, superiority, and equivalence tests, which you request with the NONINF, SUP, and EQUIV *binomial-options*, respectively. If you do not specify MARGIN=, PROC FREQ uses a margin of 0.2 by default.

For noninferiority and superiority tests, specify a single *value* for the MARGIN= option. The MARGIN= *value* must be a positive number. You can specify *value* as a number between 0 and 1. Or you can specify *value* in percentage form as a number between 1 and 100, and PROC FREQ converts that number to a proportion. The procedure treats the value 1 as 1%.

For noninferiority and superiority tests, the test limits must be between 0 and 1. The limits are determined by the null proportion value (which you can specify with the P= *binomial-option*)

and by the margin value. The noninferiority limit equals the null proportion minus the margin. By default, the null proportion equals 0.5 and the margin equals 0.2, which gives a noninferiority limit of 0.3. The superiority limit equals the null proportion plus the margin, which is 0.7 by default.

For an equivalence test, you can specify a single MARGIN= *value*, or you can specify both *lower* and *upper* values. If you specify a single MARGIN= *value*, it must be a positive number, as described previously. If you specify a single MARGIN= *value* for an equivalence test, PROC FREQ uses –*value* as the lower margin and *value* as the upper margin for the test. If you specify both *lower* and *upper* values for an equivalence test, you can specify them in proportion form as numbers between –1 or 1. Or you can specify them in percentage form as numbers between –100 and 100, and PROC FREQ converts the numbers to proportions. The value of *lower* must be less than the value of *upper*.

The equivalence limits must be between 0 and 1. The equivalence limits are determined by the null proportion value (which you can specify with the P= *binomial-option*) and by the margin values. The lower equivalence limit equals the null proportion plus the lower margin. By default, the null proportion equals 0.5 and the lower margin equals –0.2, which gives a lower equivalence limit of 0.3. The upper equivalence limit equals the null proportion plus the upper margin, which is 0.7 by default.

See the sections "Noninferiority Test" on page 2333 and "Equivalence Test" on page 2335 for details.

NONINF | NONINFERIORITY

requests a test of noninferiority for the binomial proportion. See the section "Noninferiority Test" on page 2333 for details. You can specify the noninferiority test margin, the null proportion, and the variance type with the MARGIN=, P=, and VAR= *binomial-options*, respectively.

P=*value*

specifies the null hypothesis proportion for the binomial tests. If you omit the P= option, PROC FREQ uses 0.5 as the null proportion. The null proportion *value* must be a positive number. You can specify *value* as a number between 0 and 1. Or you can specify *value* in percentage form as a number between 1 and 100, and PROC FREQ converts that number to a proportion. The procedure treats the value 1 as 1%.

SUP | SUPERIORITY

requests a test of superiority for the binomial proportion. See the section "Superiority Test" on page 2334 for details. You can specify the superiority test margin, the null proportion, and the variance type with the MARGIN=, P=, and VAR= *binomial-options*, respectively.

VAR=SAMPLE | NULL

specifies the type of variance estimate to use in the Wald tests of noninferiority, superiority, and equivalence. The default is VAR=SAMPLE, which estimates the variance from the sample proportion. VAR=NULL uses a test-based variance that is computed from the null hypothesis proportion (which is specified by the P= *binomial-option*). See the sections "Noninferiority Test" on page 2333 and "Equivalence Test" on page 2335 for details.

WALD

requests Wald confidence limits for the binomial proportion. See the section "Wald Confidence Limits" on page 2330 for details. If you specify the CORRECT *binomial-option*, the Wald confidence limits include a continuity correction. If you do not request any binomial confidence limits by specifying *binomial-options*, PROC FREQ produces Wald and exact (Clopper-Pearson) confidence limits by default.

WILSON | W | SCORE

requests Wilson confidence limits for the binomial proportion. These are also known as *score* confidence limits. See the section "Wilson (Score) Confidence Limits" on page 2331 for details.

BINOMIALC < (*binomial-options*) >

requests the BINOMIAL statistics for one-way tables and includes a continuity correction in the Wald confidence limits and tests. Specifying BINOMIALC is equivalent to specifying the BINOMIAL(CORRECT) option.

The BINOMIAL statistics include the binomial proportion, its asymptotic standard error, Wald and exact (Clopper-Pearson) confidence limits, and the asymptotic equality test for the binomial proportion by default. You can request exact binomial tests by specifying the BINOMIAL option in the EXACT statement.

You can specify *binomial-options* inside parentheses following BINOMIALC to request additional tests and confidence limits for the binomial proportion. The *binomial-options* are the same as those available with the BINOMIAL option (Table 36.10). See the description of the BINOMIAL option and the section "Binomial Proportion" on page 2329 for details.

CELLCHI2

displays each crosstabulation table cell's contribution to the total Pearson chi-square statistic. The cell contribution is computed as

$$\frac{(frequency - expected)^2}{expected}$$

where *frequency* is the table cell frequency or count and *expected* is the expected cell frequency, which is computed under the null hypothesis that the row and column variables are independent. See the section "Pearson Chi-Square Test for Two-Way Tables" on page 2316 for details.

The CELLCHI2 option has no effect for one-way tables or for tables that are displayed with the LIST option.

CHISQ < (*option*) >

requests chi-square tests of homogeneity or independence and measures of association based on the chi-square statistic. The tests include the Pearson chi-square, likelihood-ratio chi-square, and Mantel-Haenszel chi-square. The measures include the phi coefficient, the contingency coefficient, and Cramer's V. For 2×2 tables, the CHISQ option also provides Fisher's exact test and the continuity-adjusted chi-square. See the section "Chi-Square Tests and Statistics" on page 2315 for details.

For one-way tables, the CHISQ option provides a chi-square goodness-of-fit test for equal proportions. If you specify the null hypothesis proportions with the TESTP= option, PROC FREQ computes a chi-square goodness-of-fit test for the specified proportions. If you specify null hypothesis frequencies with the TESTF= option, PROC FREQ computes a chi-square goodness-of-fit test for the

specified frequencies. See the section "Chi-Square Test for One-Way Tables" on page 2316 for more information.

To request Fisher's exact test for tables larger than 2×2, use the FISHER option in the EXACT statement. Exact tests are also available for other CHISQ statistics, including the Pearson, likelihood-ratio, and Mantel-Haenszel chi-square, and the chi-square goodness-of-fit test for one-way tables. You can use the EXACT statement to request these tests. See the section "Exact Statistics" on page 2366 for details.

You can specify the following *option* in parentheses following the CHISQ option:

WARN=*value* | **(***values***)**
controls the warning message about the validity of the asymptotic Pearson chi-square test. By default, PROC FREQ displays a warning when more than 20% of the table cells have expected frequencies that are less than 5. If you specify the NOPRINT option in the PROC FREQ statement, this warning is included in the log; otherwise, the warning is displayed as a footnote in the chi-square table. You can use the WARN= option to suppress the warning and to include a warning indicator in the output data set.

The WARN= option can take one or more of the following values. If you specify more than one value, enclose the values in parentheses following WARN=. For example, `warn = (output noprint)`.

Value of WARN=	Description
OUTPUT	Adds a warning indicator variable to the output data set
NOLOG	Suppresses the chi-square warning message in the log
NOPRINT	Suppresses the chi-square warning message in the display
NONE	Suppresses the chi-square warning message entirely

If you specify the WARN=OUTPUT option, the chi-square ODS output data set contains a variable named Warning that equals 1 for the Pearson chi-square when more than 20% of the table cells have expected frequencies that are less than 5 and equals 0 otherwise. If you specify WARN=OUTPUT and also specify the CHISQ option in the OUTPUT statement, then the statistics output data set contains a variable named WARN_PCHI that indicates the warning.

The WARN=NOLOG option has the same effect as the NOWARN option in the TABLES statement.

CL
requests confidence limits for the MEASURES statistics. If you omit the MEASURES option, the CL option invokes MEASURES. You can set the level of the confidence limits by using the ALPHA= option. The default of ALPHA=0.5 produces 95% confidence limits. See the sections "Measures of Association" on page 2320 and "Confidence Limits" on page 2320 for more information.

CMH <(*cmh-options***)>**
requests Cochran-Mantel-Haenszel statistics, which test for association between the row and column variables after adjusting for the remaining variables in a multiway table. The Cochran-Mantel-Haenszel statistics include the nonzero correlation statistic, the row mean scores (ANOVA) statistic,

and the general association statistic. In addition, for 2×2 tables, the CMH option provides the adjusted Mantel-Haenszel and logit estimates of the odds ratio and relative risks, together with their confidence limits. For stratified 2 × 2 tables, the CMH option provides the Breslow-Day test for homogeneity of odds ratios. (To request Tarone's adjustment for the Breslow-Day test, specify the BDT *cmh-option*.) See the section "Cochran-Mantel-Haenszel Statistics" on page 2357 for details.

You can use the CMH1 or CMH2 option to control the number of CMH statistics that PROC FREQ computes.

For stratified 2 × 2 tables, you can request Zelen's exact test for equal odds ratios by specifying the EQOR option in the EXACT statement. See the section "Zelen's Exact Test for Equal Odds Ratios" on page 2363 for details. You can request exact confidence limits for the common odds ratio by specifying the COMOR option in the EXACT statement. This option also provides a common odds ratio test. See the section "Exact Confidence Limits for the Common Odds Ratio" on page 2363 for details.

You can specify the following *cmh-options* in parentheses following the CMH option. These *cmh-options*, which apply to stratified 2 × 2 tables, are also available with the CMH1 or CMH2 option.

BDT

requests Tarone's adjustment in the Breslow-Day test for homogeneity of odds ratios. See the section "Breslow-Day Test for Homogeneity of the Odds Ratios" on page 2363 for details. The BDT *cmh-option* has the same effect as the BDT option in the TABLES statement.

GAILSIMON | GS <(COLUMN=1 | 2)>

requests the Gail-Simon test for qualitative interaction, which applies to stratified 2 × 2 tables. See the section "Gail-Simon Test for Qualitative Interactions" on page 2365 for details.

The COLUMN= option specifies the column of the risk differences to use in computing the Gail-Simon test. By default, PROC FREQ uses column 1 risk differences. If you specify COLUMN=2, PROC FREQ uses column 2 risk differences.

The GAILSIMON *cmh-option* has the same effect as the GAILSIMON option in the TABLES statement.

MANTELFLEISS | MF

requests the Mantel-Fleiss criterion for the Mantel-Haenszel statistic for stratified 2 × 2 tables. See the section "Mantel-Fleiss Criterion" on page 2360 for details.

CMH1 <(*cmh-options*)>

requests the Cochran-Mantel-Haenszel correlation statistic. This option does not provide the CMH row mean scores (ANOVA) statistic or the general association statistic, which are provided by the CMH option. For tables larger than 2 × 2, the CMH1 option requires less memory than the CMH option, which can require an enormous amount of memory for large tables.

For 2 × 2 tables, the CMH1 option also provides the adjusted Mantel-Haenszel and logit estimates of the odds ratio and relative risks, together with their confidence limits. For stratified 2 × 2 tables, the CMH1 option provides the Breslow-Day test for homogeneity of odds ratios.

The *cmh-options* available with the CMH1 option are the same as those available with the CMH option. See the description of the CMH option for details.

CMH2 <(*cmh-options*)>

requests the Cochran-Mantel-Haenszel correlation statistic and the row mean scores (ANOVA) statistic. This option does not provide the CMH general association statistic, which is provided by the CMH option. For tables larger than 2×2, the CMH2 option requires less memory than the CMH option, which can require an enormous amount of memory for large tables.

For 2×2 tables, the CMH1 option also provides the adjusted Mantel-Haenszel and logit estimates of the odds ratio and relative risks, together with their confidence limits. For stratified 2×2 tables, the CMH1 option provides the Breslow-Day test for homogeneity of odds ratios.

The *cmh-options* available with the CMH2 option are the same as those available with the CMH option. See the description of the CMH option for details.

CONTENTS=*'string'*

specifies the label to use for crosstabulation tables in the contents file, the Results window, and the trace record. For information about output presentation, see the *SAS Output Delivery System: User's Guide*.

If you omit the CONTENTS= option, the contents label for crosstabulation tables is "Cross-Tabular Freq Table" by default.

Note that contents labels for all crosstabulation tables that are produced by a single TABLES statement use the same text. To specify different contents labels for different crosstabulation tables, request the tables in separate TABLES statements and use the CONTENTS= option in each TABLES statement.

To remove the crosstabulation table entry from the contents file, you can specify a null label with CONTENTS=''.

The CONTENTS= option affects only contents labels for crosstabulation tables. It does not affect contents labels for other PROC FREQ tables.

To specify the contents label for any PROC FREQ table, you can use PROC TEMPLATE to create a customized table definition. The CONTENTS_LABEL attribute in the DEFINE TABLE statement of PROC TEMPLATE specifies the contents label for the table. See the chapter "The TEMPLATE Procedure" in the *SAS Output Delivery System: User's Guide* for more information.

CONVERGE=*value*

specifies the convergence criterion for computing the polychoric correlation, which you request with the PLCORR option. The CONVERGE= *value* must be a positive number. By default CONVERGE=0.0001. Iterative computation of the polychoric correlation stops when the convergence measure falls below the value of CONVERGE= or when the number of iterations exceeds the value specified in the MAXITER= option, whichever happens first. See the section "Polychoric Correlation" on page 2326 for details.

CROSSLIST

displays crosstabulation tables in ODS column format instead of the default crosstabulation cell format. In a CROSSLIST table display, the rows correspond to the crosstabulation table cells, and the columns correspond to descriptive statistics such as Frequency and Percent. The CROSSLIST table displays the same information as the default crosstabulation table, but uses an ODS column format instead of the table cell format. See the section "Multiway Tables" on page 2376 for details about the contents of the CROSSLIST table.

You can control the contents of a CROSSLIST table with the same options available for the default crosstabulation table. These include the NOFREQ, NOPERCENT, NOROW, and NOCOL options. You can request additional information in a CROSSLIST table with the CELLCHI2, DEVIATION, EXPECTED, MISSPRINT, and TOTPCT options.

The FORMAT= option and the CUMCOL option have no effect for CROSSLIST tables. You cannot specify both the LIST option and the CROSSLIST option in the same TABLES statement.

You can use the NOSPARSE option to suppress display of variable levels with zero frequency in CROSSLIST tables. By default for CROSSLIST tables, PROC FREQ displays all levels of the column variable within each level of the row variable, including any column variable levels with zero frequency for that row. And for multiway tables displayed with the CROSSLIST option, the procedure displays all levels of the row variable for each stratum of the table by default, including any row variable levels with zero frequency for the stratum.

CUMCOL

displays the cumulative column percentages in the cells of the crosstabulation table. The CUMCOL option does not apply to crosstabulation tables produced with the LIST or CROSSLIST option.

DEVIATION

displays the deviation of the frequency from the expected frequency for each cell of the crosstabulation table. See the section "Pearson Chi-Square Test for Two-Way Tables" on page 2316 for details. The DEVIATION option does not apply to crosstabulation tables produced with the LIST option.

EXPECTED

displays the expected cell frequencies under the hypothesis of independence (or homogeneity) for crosstabulation tables. See the section "Pearson Chi-Square Test for Two-Way Tables" on page 2316 for details. The EXPECTED option does not apply to tables produced with the LIST option.

FISHER | EXACT

requests Fisher's exact test for tables that are larger than 2×2. (For 2×2 tables, the CHISQ option provides Fisher's exact test.) This test is also known as the Freeman-Halton test. See the sections "Fisher's Exact Test" on page 2318 and "Exact Statistics" on page 2366 for more information.

If you omit the CHISQ option in the TABLES statement, the FISHER option invokes CHISQ. You can also request Fisher's exact test by specifying the FISHER option in the EXACT statement.

NOTE: PROC FREQ computes exact tests with fast and efficient algorithms that are superior to direct enumeration. Exact tests are appropriate when a data set is small, sparse, skewed, or heavily tied. For some large problems, computation of exact tests might require a considerable amount of time and memory. Consider using asymptotic tests for such problems. Alternatively, when asymptotic methods might not be sufficient for such large problems, consider using Monte Carlo estimation of exact *p*-values. See the section "Computational Resources" on page 2369 for more information.

FORMAT=*format-name*

specifies a format for the following crosstabulation table cell values: frequency, expected frequency, and deviation. PROC FREQ also uses the specified format to display the row and column total frequencies and the overall total frequency in crosstabulation tables.

You can specify any standard SAS numeric format or a numeric format defined with the FORMAT procedure. The format length must not exceed 24. If you omit the FORMAT= option, by default

PROC FREQ uses the BEST6. format to display frequencies less than 1E6, and the BEST7. format otherwise.

The FORMAT= option applies only to crosstabulation tables displayed in the default format. It does not apply to crosstabulation tables produced with the LIST or CROSSLIST option.

To change display formats in any FREQ table, you can use PROC TEMPLATE. See the chapter "The TEMPLATE Procedure" in the *SAS Output Delivery System: User's Guide* for more information.

GAILSIMON | GS <(COLUMN=1 | 2)>

requests the Gail-Simon test for qualitative interaction, which applies to stratified 2×2 tables. See the section "Gail-Simon Test for Qualitative Interactions" on page 2365 for details.

The COLUMN= option specifies the column of the risk differences to use in computing the Gail-Simon test. By default, PROC FREQ uses column 1 risk differences. If you specify COLUMN=2, PROC FREQ uses column 2 risk differences.

JT

requests the Jonckheere-Terpstra test. See the section "Jonckheere-Terpstra Test" on page 2350 for details.

LIST

displays two-way to n-way crosstabulation tables in a list format instead of the default crosstabulation cell format. The LIST option displays the entire multiway table in one table, instead of displaying a separate two-way table for each stratum. See the section "Multiway Tables" on page 2376 for details.

The LIST option is not available when you also specify statistical options. You must use the standard crosstabulation table display or the CROSSLIST display when you request statistical tests or measures.

MAXITER=*number*

specifies the maximum number of iterations for computing the polychoric correlation, which you request with the PLCORR option. The value of the MAXITER= option must be a positive integer. By default MAXITER=20. Iterative computation of the polychoric correlation stops when the number of iterations exceeds the MAXITER= value or when the convergence measures falls below the value of the CONVERGE= option, whichever happens first. See the section "Polychoric Correlation" on page 2326 for details.

MEASURES

requests several measures of association and their asymptotic standard errors. The MEASURES option provides the following statistics: gamma, Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$, Somers' $D(R|C)$, the Pearson and Spearman correlation coefficients, lambda (symmetric and asymmetric), and uncertainty coefficients (symmetric and asymmetric). To request confidence limits for these measures of association, you can specify the CL option.

For 2×2 tables, the MEASURES option also provides the odds ratio, column 1 relative risk, column 2 relative risk, and the corresponding confidence limits. Alternatively, you can obtain the odds ratio and relative risks, without the other measures of association, by specifying the RELRISK option.

See the section "Measures of Association" on page 2320 for details.

You can use the TEST statement to request asymptotic tests for the following measures of association: gamma, Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$, Somers' $D(R|C)$, and the Pearson and

Spearman correlation coefficients. You can use the EXACT statement to request exact confidence limits for the odds ratio, exact unconditional confidence limits for the relative risks, and exact tests for the following measures of association: Kendall's tau-*b*, Stuart's tau-*c*, Somers' $D(C|R)$ and $D(R|C)$, and the Pearson and Spearman correlation coefficients. See the section "Exact Statistics" on page 2366 for more information.

MISSING

treats missing values as a valid nonmissing level for all TABLES variables. The MISSING option displays the missing levels in frequency and crosstabulation tables and includes them in all calculations of percentages, tests, and measures.

By default, if you do not specify the MISSING or MISSPRINT option, an observation is excluded from a table if it has a missing value for any of the variables in the TABLES request. When PROC FREQ excludes observations with missing values, it displays the total frequency of missing observations below the table. See the section "Missing Values" on page 2310 for more information.

MISSPRINT

displays missing value frequencies in frequency and crosstabulation tables but does not include the missing value frequencies in any computations of percentages, tests, or measures.

By default, if you do not specify the MISSING or MISSPRINT option, an observation is excluded from a table if it has a missing value for any of the variables in the TABLES request. When PROC FREQ excludes observations with missing values, it displays the total frequency of missing observations below the table. See the section "Missing Values" on page 2310 for more information.

NOCOL

suppresses the display of column percentages in crosstabulation table cells.

NOCUM

suppresses the display of cumulative frequencies and percentages in one-way frequency tables. The NOCUM option also suppresses the display of cumulative frequencies and percentages in crosstabulation tables in list format, which you request with the LIST option.

NOFREQ

suppresses the display of cell frequencies in crosstabulation tables. The NOFREQ option also suppresses row total frequencies. This option has no effect for one-way tables or for crosstabulation tables in list format, which you request with the LIST option.

NOPERCENT

suppresses the display of overall percentages in crosstabulation tables. These percentages include the cell percentages of the total (two-way) table frequency, as well as the row and column percentages of the total table frequency. To suppress the display of cell percentages of row or column totals, use the NOROW or NOCOL option, respectively.

For one-way frequency tables and crosstabulation tables in list format, the NOPERCENT option suppresses the display of percentages and cumulative percentages.

NOPRINT

suppresses the display of frequency and crosstabulation tables but displays all requested tests and statistics. To suppress the display of all output, including tests and statistics, use the NOPRINT option in the PROC FREQ statement.

NOROW

suppresses the display of row percentages in crosstabulation table cells.

NOSPARSE

suppresses the display of cells with a zero frequency count in LIST output and omits them from the OUT= data set. The NOSPARSE option applies when you specify the ZEROS option in the WEIGHT statement to include observations with zero weights. By default, the ZEROS option invokes the SPARSE option, which displays table cells with a zero frequency count in the LIST output and includes them in the OUT= data set. See the description of the ZEROS option for more information.

The NOSPARSE option also suppresses the display of variable levels with zero frequency in CROSSLIST tables. By default for CROSSLIST tables, PROC FREQ displays all levels of the column variable within each level of the row variable, including any column variable levels with zero frequency for that row. For multiway tables displayed with the CROSSLIST option, the procedure displays all levels of the row variable for each stratum of the table by default, including any row variable levels with zero frequency for the stratum.

NOWARN

suppresses the log warning message about the validity of the asymptotic Pearson chi-square test. By default, PROC FREQ provides a warning about the validity of the asymptotic Pearson chi-square test when more than 20cells have expected frequencies that are less than 5. This warning message appears in the log if you specify the NOPRINT option in the PROC FREQ statement,

The NOWARN option is equivalent to the CHISQ(WARN=NOLOG) option. You can also use the CHISQ(WARN=) option to suppress the warning message in the display and to request a warning variable in the chi-square ODS output data set or in the OUTPUT data set.

OUT=*SAS-data-set*

names an output data set that contains frequency or crosstabulation table counts and percentages. If more than one table request appears in the TABLES statement, the contents of the OUT= data set correspond to the last table request in the TABLES statement. The OUT= data set variable COUNT contains the frequencies and the variable PERCENT contains the percentages. See the section "Output Data Sets" on page 2371 for details. You can specify the following options to include additional information in the OUT= data set: OUTCUM, OUTEXPECT, and OUTPCT.

OUTCUM

includes cumulative frequencies and cumulative percentages in the OUT= data set for one-way tables. The variable CUM_FREQ contains the cumulative frequencies, and the variable CUM_PCT contains the cumulative percentages. See the section "Output Data Sets" on page 2371 for details. The OUTCUM option has no effect for two-way or multiway tables.

OUTEXPECT

includes expected cell frequencies in the OUT= data set for crosstabulation tables. The variable EXPECTED contains the expected cell frequencies. See the section "Output Data Sets" on page 2371 for details. The EXPECTED option has no effect for one-way tables.

OUTPCT

includes the following additional variables in the OUT= data set for crosstabulation tables:

PCT_COL	percentage of column frequency
PCT_ROW	percentage of row frequency
PCT_TABL	percentage of stratum (two-way table) frequency, for n-way tables where $n > 2$

See the section "Output Data Sets" on page 2371 for details. The OUTPCT option has no effect for one-way tables.

PLCORR

requests the polychoric correlation coefficient. For 2×2 tables, this statistic is more commonly known as the tetrachoric correlation coefficient, and it is labeled as such in the displayed output. See the section "Polychoric Correlation" on page 2326 for details. Also see the descriptions of the CONVERGE= and MAXITER= options, which you can specify to control the iterative computation of the polychoric correlation coefficient.

If you omit the MEASURES option, the PLCORR option invokes MEASURES.

PLOTS < (*global-plot-options*) > < = *plot-request* < (*plot-options*) > >

PLOTS < (*global-plot-options*) >

< = (*plot-request* < (*plot-options*) > < ... *plot-request* < (*plot-options*) > >) >

controls the plots that are produced through ODS Graphics. *Plot-requests* identify the plots, and *plot-options* control the appearance and content of the plots. You can specify *plot-options* in parentheses following a *plot-request*. A *global-plot-option* applies to all plots for which it is available, unless it is altered by a specific *plot-option*. You can specify *global-plot-options* in parentheses following the PLOTS option.

When you specify only one *plot-request*, you can omit the parentheses around the request. For example:

```
plots=all
plots=freqplot
plots=(freqplot oddsratioplot)
plots(only)=(cumfreqplot deviationplot)
```

ODS Graphics must be enabled before requesting plots. For example:

```
ods graphics on;
proc freq;
   tables treatment*response / chisq plots=freqplot;
   weight wt;
run;
ods graphics off;
```

For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 609 in Chapter 21, "Statistical Graphics Using ODS."

If ODS Graphics is enabled but you do not specify the PLOTS= option, PROC FREQ produces all plots that are associated with the analyses that you request with the exception of the frequency and cumulative frequency plots. To produce a frequency plot or cumulative frequency plot when ODS Graphics is enabled, you must specify the FREQPLOT or CUMFREQPLOT *plot-request*, respectively, in the PLOTS= option. PROC FREQ produces the remaining plots (listed in Table 36.11) by default when you request the corresponding TABLES statement options. You can suppress default plots and request specific plots by using the PLOTS(ONLY)= option; PLOTS(ONLY)=(*plot-requests*) produces only the plots that are specified as *plot-requests*. You can suppress all plots with the PLOTS=NONE option. The PLOTS= option has no effect when you specify the NOPRINT option in the PROC FREQ statement.

Table 36.11 lists the available *plot-requests*, together with their *plot-options* and required TABLES statement options.

Table 36.11 PLOTS= Options

Plot Request	Plot Options	Required TABLES Statement Option
AGREEPLOT	LEGEND= PARTIAL= SHOWSCALE= STATS	AGREE ($r \times r$ table)
CUMFREQPLOT	ORIENT= SCALE= TYPE=	One-way table request
DEVIATIONPLOT	NOSTAT ORIENT= TYPE=	CHISQ (one-way table)
FREQPLOT	ORIENT= SCALE= TYPE=	Any table request
FREQPLOT	NPANELPOS= TWOWAY=	Two-way or multiway table request
KAPPAPLOT	CLDISPLAY= NPANELPOS= ORDER= RANGE= STATS	AGREE ($h \times r \times r$ table)
ODDSRATIOPLOT	CLDISPLAY= EXACT* LOGBASE= NPANELPOS= ORDER= RANGE= STATS	MEASURES or RELRISK ($h \times 2 \times 2$ table)

* Also requires EXACT statement

Table 36.11 *continued*

Plot Request	Plot Options	Required TABLES Statement Option
RELRISKPLOT	CLDISPLAY= COLUMN= EXACT* LOGBASE= NPANELPOS= ORDER= RANGE= STATS	MEASURES or RELRISK ($h \times 2 \times 2$ table)
RISKDIFFPLOT	CLDISPLAY= COLUMN= EXACT* NPANELPOS= ORDER= RANGE= STATS	RISKDIFF ($h \times 2 \times 2$ table)
WTKAPPAPLOT	CLDISPLAY= NPANELPOS= ORDER= RANGE= STATS	AGREE ($h \times r \times r$ table, $r > 2$)

Global Plot Options

A *global-plot-option* applies to all plots for which the option is available, unless it is altered by a specific *plot-option*. You can specify *global-plot-options* in parentheses following the PLOTS option.

The following specific *plot-options* are available as *global-plot-options*: CLDISPLAY=, COLUMN=, EXACT, LOGBASE=, NPANELPOS=, ORDER=, ORIENT=, RANGE=, SCALE=, STATS, and TYPE=.

These *plot-options* are described in the section "Plot Options." Additionally, you can specify the following *global-plot-option* in parentheses following the PLOTS option:

ONLY
suppresses the default plots and requests only the plots that are specified as *plot-requests*.

Plot Requests

The following *plot-requests* are available with the PLOTS= option:

AGREEPLOT
requests an agreement plot (Bangdiwala and Bryan 1987). An agreement plot displays the strength of agreement in a two-way table, where the row and column variables represent two

independent ratings of n subjects. See Bangdiwala (1988), Bangdiwala et al. (2008), and Friendly (2000, Section 3.7.2) for information about agreement plots.

To produce an agreement plot, you must also specify the AGREE option in the TABLES statement. Agreement statistics and plots are available for two-way square tables, where the number of rows equals the number of columns. The following *plot-options* are available for AGREEPLOT: LEGEND=, PARTIAL=, SHOWSCALE=, and STATS.

ALL

requests all plots that are associated with the specified analyses. This is the default if you do not specify the PLOTS(ONLY) option.

CUMFREQPLOT < (*plot-options*) >

requests a plot of cumulative frequencies. Cumulative frequency plots are available for one-way frequency tables. The following *plot-options* are available for CUMFREQPLOT: ORIENT=, SCALE=, and TYPE=.

You must specify the CUMFREQPLOT *plot-request* in the PLOTS= option to produce a cumulative frequency plot. Cumulative frequency plots are not produced by default when you request frequency or crosstabulation tables.

DEVIATIONPLOT < (*plot-options*) >

requests a plot of relative deviations from expected frequencies. Deviation plots are available for one-way frequency tables. To produce a deviation plot, you must also specify the CHISQ option in the TABLES statement. The following *plot-options* are available for DEVIATIONPLOT: NOSTAT, ORIENT=, and TYPE=.

FREQPLOT < (*plot-options*) >

requests a frequency plot. Frequency plots are available for frequency and crosstabulation tables. For multiway tables, PROC FREQ provides a two-way frequency plot for each stratum.

The following *plot-options* are available for FREQPLOT for all tables: ORIENT=, SCALE=, and TYPE=. Additionally, the TWOWAY= and NPANELPOS= *plot-options* are available for two-way and multiway tables. You can use the TWOWAY= *plot-option* to specify the layout of a two-way frequency plot. The NPANELPOS= *plot-option* is not available with the TWOWAY=STACKED layout.

You must specify the FREQPLOT *plot-request* in the PLOTS= option to produce a frequency plot. Frequency plots are not produced by default when you request frequency or crosstabulation tables.

KAPPAPLOT < (*plot-options*) >

requests a plot of kappa statistics and confidence limits. Kappa plots are available for multiway square tables. To produce a kappa plot, you must also specify the AGREE option in the TABLES statement. The following *plot-options* are available for KAPPAPLOT: CLDISPLAY=, NPANELPOS=, ORDER=, RANGE=, and STATS.

NONE

suppresses all plots.

ODDSRATIOPLOT < (*plot-options*) >
requests a plot of odds ratios with confidence limits. Odds ratio plots are available for multiway 2×2 tables. To produce an odds ratio plot, you must also specify the MEASURES or RELRISK option in the TABLES statement. The following *plot-options* are available for ODDSRATIOPLOT: CLDISPLAY=, EXACT, LOGBASE=, NPANELPOS=, ORDER=, RANGE=, and STATS. If you request a plot of exact confidence limits by specifying the EXACT *plot-option*, you must also request computation of exact confidence limits by specifying the OR option in the EXACT statement.

RELRISKPLOT < (*plot-options*) >
requests a plot of relative risks with confidence limits. Relative risk plots are available for multiway 2×2 tables. To produce a relative risk plot, you must also specify the MEASURES or RELRISK option in the TABLES statement. The following *plot-options* are available for RELRISKPLOT: CLDISPLAY=, COLUMN=, EXACT, LOGBASE=, NPANELPOS=, ORDER=, RANGE=, and STATS. If you request a plot of exact confidence limits by specifying the EXACT *plot-option*, you must also request computation of exact confidence limits by specifying the RELRISK option in the EXACT statement.

RISKDIFFPLOT < (*plot-options*) >
requests a plot of risk (proportion) differences with confidence limits. Risk difference plots are available for multiway 2×2 tables. To produce a risk difference plot, you must also specify the RISKDIFF option in the TABLES statement. The following *plot-options* are available for RISKDIFFPLOT: CLDISPLAY=, COLUMN=, EXACT, NPANELPOS=, ORDER=, RANGE=, and STATS.

If you request a plot of exact confidence limits by specifying the EXACT *plot-option*, you must also request computation of exact confidence limits by specifying the RISKDIFF option in the EXACT statement. If you do not specify the EXACT *plot-option*, the risk difference plot displays the Wald confidence limits that are produced by the RISKDIFF option by default and displayed in the "Risk Estimates" table.

WTKAPPAPLOT < (*plot-options*) >
requests a plot of weighted kappa statistics with confidence limits. Weighted kappa plots are available for multiway square tables. To produce a weighted kappa plot, you must also specify the AGREE option in the TABLES statement. Note that simple kappa and weighted kappa statistics are the same for 2×2 tables; therefore, the procedure does not present weighted kappa statistics for 2×2 tables. The following *plot-options* are available for WTKAPPAPLOT: CLDISPLAY=, NPANELPOS=, ORDER=, RANGE=, and STATS.

Plot Options

You can specify the following *plot-options* in parentheses after a *plot-request*:

CLDISPLAY=SERIF | LINE | BAR < *width* >
controls the appearance of the confidence limit error bars. The CLDISPLAY= *plot-option* is available for the following plots: KAPPAPLOT, ODDSRATIOPLOT, RELRISKPLOT, RISKDIFFPLOT, and WTKAPPAPLOT.

The default value is CLDISPLAY=SERIF, which displays the confidence limits as lines with serifs. CLDISPLAY=LINE displays the confidence limits as plain lines without serifs. CLDISPLAY=BAR displays the confidence limits as bars. By default, the width of the bars equals the size of the marker for the estimate. You can control the width of the bars and the size of the marker by specifying the value of *width* as a percentage of the distance between bars, $0 < \textit{width} \leq 1$. The bar might disappear when the value of *width* is very small.

COLUMN=1 | 2

specifies the 2×2 table column to use to compute the risk (proportion). The COLUMN= *plot-option* is available for the relative risk plot (RELRISKPLOT) and the risk difference plot (RISKDIFFPLOT). If you specify COLUMN=1, the plot displays the column 1 risk differences or the column 1 relative risks. Similarly, if you specify COLUMN=2, the plot displays the column 2 risk differences or relative risks. The default is COLUMN=1.

EXACT

requests exact confidence limits instead of asymptotic confidence limits. The EXACT *plot-option* is available for the odds ratio plot (ODDSRATIOPLOT), the relative risk plot (RELRISKPLOT), and the risk difference plot (RISKDIFFPLOT). If you specify the EXACT *plot-option*, you must also request computation of the exact confidence limits by specifying the corresponding option in the EXACT statement.

LOGBASE=2 | E | 10

applies to the odds ratio plot (ODDSRATIOPLOT) and the relative risk plot (RELRISKPLOT). LOGBASE= displays the odds ratio or relative risk axis on the specified log scale.

LEGEND=YES | NO

applies to the agreement plot (AGREEPLOT). LEGEND=NO suppresses the legend that identifies the areas of exact and partial agreement. The default is LEGEND=YES.

NOSTAT

applies to the deviation plot (DEVIATIONPLOT). NOSTAT suppresses the chi-square p-value that is displayed by default in the deviation plot.

NPANELPOS=*n*

applies to the following plots: FREQPLOT (for two-way and multiway tables), KAPPAPLOT, ODDSRATIOPLOT, RELRISKPLOT, RISKDIFFPLOT, and WTKAPPAPLOT.

NPANELPOS= divides the plot into multiple panels that display at most $|n|$ statistics per panel. If n is positive, the number of statistics per panel is balanced; but if n is negative, the number of statistics per panel is not balanced. By default, $n = 0$ and all statistics are displayed in a single plot. For example, suppose you want to display 21 odds ratios. Then NPANELPOS=20 displays two panels, the first with 11 odds ratios and the second with 10; NPANELPOS=–20 displays 20 odds ratios in the first panel but only one in the second.

For two-way frequency plots, NPANELPOS= divides the plot into multiple panels that display at most $|n|$ levels of the row variable per panel. The NPANELPOS= *plot-option* applies to two-way plots that are displayed with grouped layout, which you specify with the TWOWAY=GROUPVERTICAL or TWOWAY=GROUPHORIZONTAL *plot-option*. The NPANELPOS= *plot-option* does not apply to the TWOWAY=STACKED layout.

ORDER=ASCENDING | DESCENDING

displays the statistics in sorted order. By default, the statistics are displayed in the order in which the corresponding strata appear in the multiway table display. The ORDER= *plot-option* applies to the following plots: KAPPAPLOT, ODDSRATIOPLOT, RELRISKPLOT, RISKDIFFPLOT, and WTKAPPAPLOT.

ORIENT=HORIZONTAL | VERTICAL

controls the orientation of the plot. The ORIENT= *plot-option* applies to the following plots: CUMFREQPLOT, DEVIATIONPLOT, and FREQPLOT.

ORIENT=HORIZONTAL places the variable levels on the Y axis and the frequencies, percentages, or statistic-values on the X axis. ORIENT=VERTICAL places the variable levels on the X axis. The default orientation is ORIENT=VERTICAL for bar charts (TYPE=BARCHART) and ORIENT=HORIZONTAL for dot plots (TYPE=DOTPLOT).

PARTIAL=YES | NO

controls the display of partial agreement in the agreement plot (AGREEPLOT). PARTIAL=NO suppresses the display of partial agreement. When you specify PARTIAL=NO, the agreement plot displays only exact agreement. Exact agreement includes the diagonal cells of the square table, where the row and column variable levels are the same. Partial agreement includes the adjacent off-diagonal table cells, where the row and column values are within one level of exact agreement. The default is PARTIAL=YES.

RANGE=(< *min* **> <** *,max* **>) | CLIP**

specifies the range of values to display. The RANGE= *plot-option* applies to the following plots: KAPPAPLOT, ODDSRATIOPLOT, RELRISKPLOT, RISKDIFFPLOT, and WTKAPPAPLOT. If you specify RANGE=CLIP, the confidence limits are clipped and the display range is determined by the minimum and maximum values of the estimates. By default, the display range includes all confidence limits.

SCALE=FREQ | LOG | PERCENT | SQRT

specifies the scale of the frequencies to display. The SCALE= *plot-option* applies to the frequency plot (FREQPLOT) and the cumulative frequency plot (CUMFREQPLOT).

The default is SCALE=FREQ, which displays unscaled frequencies. SCALE=LOG displays log (base 10) frequencies. SCALE=PERCENT displays percentages (relative frequencies). SCALE=SQRT displays square roots of the frequencies, which produces a plot known as a *rootogram*.

SHOWSCALE=YES | NO

controls the display of the cumulative frequency scale on the right side of the agreement plot (AGREEPLOT). SHOWSCALE=NO suppresses the display of the scale. The default is SHOWSCALE=YES.

STATS

displays statistic values in the plot. For the following plots, the STATS *plot-option* displays the statistics and their confidence limits on the right side of the plot: KAPPAPLOT, ODDSRATIOPLOT, RELRISKPLOT, RISKDIFFPLOT, and WTKAPPAPLOT.

For the agreement plot (AGREEPLOT), STATS displays the values of the kappa statistic, the weighted kappa statistic, and the B_n measure (Bangdiwala 1987).

If you do not request the STATS *plot-option*, these plots do not display the statistic values.

TWOWAY=GROUPVERTICAL | GROUPHORIZONTAL | STACKED

specifies the layout for a two-way frequency plot (FREQPLOT). The TWOWAY= *plot-option* applies to frequency plots for two-way and multiway table requests; PROC FREQ produces a two-way frequency plot for each stratum of a multiway table request.

TWOWAY=GROUPVERTICAL produces a grouped plot with a vertical common baseline. The plot is grouped by the row variable, which is the first variable that you specify in a two-way table request. TWOWAY=GROUPHORIZONTAL produces a grouped plot with a horizontal common baseline.

TWOWAY=STACKED produces stacked frequency plots for two-way tables. In a stacked bar chart, the bars correspond to the column variable values, and the row frequencies are stacked within each column. In a stacked dot plot, the dotted lines correspond to the columns, and the row frequencies within columns are plotted as data dots on the same column line.

The default two-way layout is TWOWAY=GROUPVERTICAL. The ORIENT=, SCALE=, and TYPE= *plot-options* are available for each TWOWAY= layout.

TYPE=BARCHART | DOTPLOT

specifies the plot type for frequency (FREQPLOT), cumulative frequency (CUMFREQPLOT), and deviation plots (DEVIATIONPLOT). TYPE=BARCHART produces a bar chart, and TYPE=DOTPLOT produces a dot plot. The default is TYPE=BARCHART.

PRINTKWT

displays the weights that PROC FREQ uses to compute the weighted kappa coefficient. You must also specify the AGREE option to request the weighted kappa coefficient. You can specify (WT=FC) with the AGREE option to request Fleiss-Cohen weights. By default, PROC FREQ uses Cicchetti-Allison weights to compute the weighted kappa coefficient. See the section "Weighted Kappa Coefficient" on page 2354 for details.

RELRISK | OR

requests relative risk measures and their confidence limits for 2 × 2 tables. These measures include the odds ratio and the column 1 and 2 relative risks. See the section "Odds Ratio and Relative Risks for 2 x 2 Tables" on page 2346 for details.

You can also obtain the RELRISK measures by specifying the MEASURES option, which produces other measures of association in addition to the relative risks.

You can request exact confidence limits for the odds ratio by specifying the OR option in the EXACT statement. You can request exact unconditional confidence limits for the relative risks by specifying the RELRISK option in the EXACT statement. See the sections "Exact Confidence Limits for the Odds Ratio" on page 2347 and "Exact Unconditional Confidence Limits for the Relative Risk" on page 2348 for more information.

RISKDIFF < (*riskdiff-options*) >

requests risks (binomial proportions) and risk differences for 2 × 2 tables. When you specify the RISKDIFF option, PROC FREQ provides the row 1 risk, row 2 risk, total (overall) risk, and risk difference (row 1 – row 2), together with their asymptotic standard errors and Wald confidence limits. PROC FREQ also provides exact (Clopper-Pearson) confidence limits for the row 1, row 2, and total risks by default. You can request exact unconditional confidence limits for the risk difference by specifying the RISKDIFF option in the EXACT statement. See the section "Risks and Risk Differences"

on page 2336 for details. PROC FREQ displays these results in the column 1 and column 2 "Risk Estimates" tables.

You can specify *riskdiff-options* inside parentheses following the RISKDIFF option to request tests and additional confidence limits for the risk difference. Table 36.12 summarizes the *riskdiff-options*.

The EQUIV, NONINF, and SUP *riskdiff-options* request tests of equivalence, noninferiority, and superiority, respectively, for the risk difference. Available test methods include Farrington-Manning, Hauck-Anderson, and Newcombe score, in addition to the Wald test.

As part of the noninferiority, superiority, and equivalence analyses, PROC FREQ provides test-based confidence limits that have a confidence coefficient of $100(1 - 2\alpha)\%$ (Schuirmann 1999). The ALPHA= option determines the confidence level, and the default of ALPHA=0.05 produces 90% confidence limits. See the sections "Noninferiority Tests" on page 2341 and "Equivalence Tests" on page 2344 for details.

The CL= *riskdiff-option* requests confidence limits for the risk difference. Available confidence limit types include exact unconditional, Farrington-Manning, Hauck-Anderson, Newcombe score, and Wald. You can request more than one type of confidence limits in the same analysis. If you specify the CORRECT *riskdiff-option*, PROC FREQ includes continuity corrections in the Newcombe and Wald confidence limits. PROC FREQ displays the confidence limits in the "Proportion (Risk) Difference Confidence Limits" table.

The ALPHA= option determines the level of the confidence limits that the CL= *riskdiff-option* provides. The default of ALPHA=0.05 produces 95% confidence limits for the risk difference.

The CL=EXACT *riskdiff-option* displays exact unconditional confidence limits in the "Proportion (Risk) Difference Confidence Limits" table. When you use CL=EXACT, you must also request computation of the exact confidence limits by specifying the RISKDIFF option in the EXACT statement.

Table 36.12 RISKDIFF (Proportion Difference) Options

Option	Description
COLUMN=1 | 2	Specifies the risk column
CORRECT	Requests continuity correction
NORISKS	Suppresses default risk tables
Request Confidence Limits	
CL=EXACT	Displays exact confidence limits
CL=FM	Requests Farrington-Manning confidence limits
CL=HA	Requests Hauck-Anderson confidence limits
CL=NEWCOMBE	Requests Newcombe confidence limits
CL=WALD	Requests Wald confidence limits
Request Tests	
EQUAL	Requests an equality test
EQUIV | EQUIVALENCE	Requests an equivalence test
NONINF | NONINFERIORITY	Requests a noninferiority test
SUP | SUPERIORITY	Requests a superiority test
MARGIN=	Specifies the test margin
METHOD=	Specifies the test method
VAR=SAMPLE | NULL	Specifies the test variance

You can specify the following *riskdiff-options* inside parentheses following the RISKDIFF option:

CL=*type* **|** **(***types***)**
requests confidence limits for the risk difference. You can specify one or more *types* of confidence limits. PROC FREQ displays the confidence limits in the "Proportion (Risk) Difference Confidence Limits" table.

The ALPHA= option determines the confidence level, and the default of ALPHA=0.05 produces 95% confidence limits for the risk difference. This differs from the test-based confidence limits that are provided with the equivalence, noninferiority, and superiority tests (EQUIV, NONINF, and SUP), which have a confidence coefficient of $100(1-2\alpha)\%$ (Schuirmann 1999).

You can specify CL= with or without requests for risk difference tests. The confidence limits produced by CL= do not depend on the tests that you request and do not use the value of the test margin (which is specified by the MARGIN= *riskdiff-option*).

You can control the risk column for the confidence limits with the COLUMN= *riskdiff-option*. If you do not specify COLUMN=, PROC FREQ provides confidence limits for the column 1 risk difference by default.

The following *types* of confidence limits are available:

EXACT
displays exact unconditional confidence limits for the risk difference in the "Proportion (Risk) Difference Confidence Limits" table. You must also request computation of the exact confidence limits by specifying the RISKDIFF option in the EXACT statement.

PROC FREQ computes the confidence limits by inverting two separate one-sided exact tests (tail method), where the tests are based on the unstandardized risk difference by default. If you specify the RISKDIFF(METHOD=FMSCORE) option in the EXACT statement, the tests are based on the Farrington-Manning score statistic. See the RISKDIFF option in the EXACT statement and the section "Exact Unconditional Confidence Limits for the Risk Difference" on page 2345 for more information.

By default, PROC FREQ also displays these exact confidence limits in the "Risk Estimates" table. You can suppress this table by specifying the NORISKS *riskdiff-option*.

FM < (NULL=*value***) >**
requests Farrington-Manning confidence limits for the risk difference. See the subsection Farrington-Manning Confidence Limits in the section "Risk Difference Confidence Limits" on page 2338 for details.

You can specify the null value of the risk difference for the Farrington-Manning computations by including NULL=*value* in parentheses following FM. The null risk difference *value* must be between –1 and 1. If you do not specify NULL=, the computations use a null risk difference of 0 by default. This differs from the Farrington-Manning confidence limits that are provided with the equivalence, noninferiority, and superiority tests, which use a null value based on the test margin (which is specified by the MARGIN= *riskdiff-option*).

HA

requests Hauck-Anderson confidence limits for the risk difference. See the subsection Hauck-Anderson Confidence Limits in the section "Risk Difference Confidence Limits" on page 2338 for details.

NEWCOMBE | SCORE | WILSON

requests Newcombe score confidence limits for the risk difference. If you specify the CORRECT *riskdiff-option*, the Newcombe confidence limits include a continuity correction. See the section "Risk Difference Confidence Limits" on page 2338 for details.

WALD <(NULL<=*value*>)>

requests Wald confidence limits for the risk difference. If you specify the CORRECT *riskdiff-option*, the Wald confidence limits include a continuity correction.

By default, the Wald confidence limits are computed by using a sample-based variance. If you specify NULL=*value* in parentheses following WALD, the confidence limit computations use a test-based variance with a null risk difference of *value*. The null *value* must be between –1 and 1. If you specify NULL but do not specify *value*, the computations use a test-based variance with a null value of 0.

See the subsection Wald Confidence Limits in the section "Risk Difference Confidence Limits" on page 2338 for details.

COLUMN=1 | 2 | BOTH

specifies the table column for which to compute the risk difference tests (EQUAL, EQUIV, NONINF, and SUP) and the risk difference confidence limits (which are requested by the CL= *riskdiff-option*).

If you do not specify COLUMN=, PROC FREQ provides the risk difference tests and confidence limits for column 1 by default. The COLUMN= option has no effect on the "Risk Estimates" table, which is produced for both column 1 and column 2. You can suppress the "Risk Estimates" table by specifying the NORISKS *riskdiff-option*.

CORRECT

includes a continuity correction in the Wald confidence limits, Wald tests, and Newcombe score confidence limits. See the section "Risks and Risk Differences" on page 2336 for details. The RISKDIFF(CORRECT) option is equivalent to the RISKDIFFC option.

EQUAL

requests a test of the null hypothesis that the risk difference equals zero. PROC FREQ provides an asymptotic Wald test of equality. If you specify the CORRECT *riskdiff-option*, the Wald test includes a continuity correction. If you specify the VAR=NULL *riskdiff-option*, the test uses the null (test-based) variance instead of the sample-based variance. See the section "Equality Test" on page 2340 for details.

EQUIV | EQUIVALENCE

requests a test of equivalence for the risk difference. See the section "Equivalence Tests" on page 2344 for details. You can specify the equivalence test margins with the MARGIN= *riskdiff-option* and the test method with the METHOD= *riskdiff-option*. PROC FREQ uses METHOD=WALD by default.

MARGIN=*value* | **(***lower*,*upper***)**

specifies the margin for the noninferiority, superiority, and equivalence tests, which you request with the NONINF, SUP, and EQUIV *riskdiff-options*, respectively. If you do not specify MARGIN=, PROC FREQ uses a margin of 0.2 by default.

For noninferiority and superiority tests, specify a single *value* for MARGIN=. The MARGIN= *value* must be a positive number. You can specify *value* as a number between 0 and 1. Or you can specify *value* in percentage form as a number between 1 and 100, and PROC FREQ converts that number to a proportion. The procedure treats the value 1 as 1%.

For an equivalence test, you can specify a single MARGIN= *value*, or you can specify both *lower* and *upper* values. If you specify a single MARGIN= *value*, it must be a positive number, as described previously. If you specify a single MARGIN= *value* for an equivalence test, PROC FREQ uses –*value* as the lower margin and *value* as the upper margin for the test. If you specify both *lower* and *upper* values for an equivalence test, you can specify them in proportion form as numbers between –1 or 1. Or you can specify them in percentage form as numbers between –100 and 100, and PROC FREQ converts the numbers to proportions. The value of *lower* must be less than the value of *upper*.

METHOD=*method*

specifies the method for the noninferiority, superiority, and equivalence analyses, which you request with the NONINF, SUP, and EQUIV *riskdiff-options*, respectively. If you do not specify the METHOD= *riskdiff-option*, PROC FREQ uses METHOD=WALD by default.

The following *methods* are available:

FM

requests Farrington-Manning tests and test-based confidence limits for the equivalence, noninferiority, and superiority analyses. See the subsection Farrington-Manning Test in the section "Noninferiority Tests" on page 2341 for details.

HA

requests Hauck-Anderson tests and confidence limits for the equivalence, noninferiority, and superiority analyses. See the subsection Hauck-Anderson Test in the section "Noninferiority Tests" on page 2341 for details.

NEWCOMBE | SCORE | WILSON

requests Newcombe score confidence limits for the equivalence, noninferiority, and superiority analyses. If you specify the CORRECT *riskdiff-option*, the Newcombe confidence limits include a continuity correction. See the subsection Newcombe Score Confidence Limits in the section "Noninferiority Tests" on page 2341 for details.

WALD

requests Wald tests and confidence limits for the equivalence, noninferiority, and superiority analyses. If you specify the CORRECT *riskdiff-option*, the Wald confidence limits include a continuity correction. If you specify the VAR=NULL *riskdiff-option*, the tests and confidence limits use the null (test-based) variance instead of the sample-based variance. See the subsection Wald Test in the section "Noninferiority Tests" on page 2341 for details.

NONINF | NONINFERIORITY
requests a test of noninferiority for the risk difference. See the section "Noninferiority Tests" on page 2341 for details. You can specify the test margin with the MARGIN= *riskdiff-option* and the test method with the METHOD= *riskdiff-option*. PROC FREQ uses METHOD=WALD by default.

NORISKS
suppresses display of the "Risk Estimates" tables, which the RISKDIFF option produces by default for column 1 and column 2. The "Risk Estimates" tables contain the risks and risk differences, together with their asymptotic standard errors, Wald confidence limits, and exact confidence limits.

SUP | SUPERIORITY
requests a test of superiority for the binomial proportion. See the section "Superiority Test" on page 2344 for details. You can specify the test margin with the MARGIN= *riskdiff-option* and the test method with the METHOD= *riskdiff-option*. PROC FREQ uses METHOD=WALD by default.

VAR=SAMPLE | NULL
specifies the type of variance estimate to use in the Wald tests of noninferiority, superiority, equivalence, and equality. The default is VAR=SAMPLE, which uses the sample-based variance. VAR=NULL uses a test-based variance that is computed from the null hypothesis risk difference value. See the sections "Equality Test" on page 2340 and "Noninferiority Tests" on page 2341 for details.

RISKDIFFC < (*riskdiff-options*) >
requests the RISKDIFF statistics for 2×2 tables and includes continuity corrections in the Wald confidence limits, Wald tests, and Newcombe confidence limits. Specifying RISKDIFFC is equivalent to specifying the RISKDIFF(CORRECT) option.

The RISKDIFF statistics include risks (binomial proportions) and risk differences for 2×2 tables. PROC FREQ provides the row 1 risk, row 2 risk, total risk, and risk difference (row 1 – row 2), together with their asymptotic standard errors and Wald confidence limits. PROC FREQ also provides exact (Clopper-Pearson) confidence limits for the row 1, row 2, and total risks by default. You can request exact unconditional confidence limits for the risk difference by specifying the RISKDIFF option in the EXACT statement.

You can specify *riskdiff-options* inside parentheses following RISKDIFFC to request tests and additional confidence limits for the risk difference. The *riskdiff-options* are the same as those available with the RISKDIFF option (Table 36.12). See the description of the RISKDIFF option and the section "Risks and Risk Differences" on page 2336 for details.

SCORES=*type*
specifies the type of row and column scores that PROC FREQ uses to compute the following statistics: Mantel-Haenszel chi-square, Pearson correlation, Cochran-Armitage test for trend, weighted kappa coefficient, and Cochran-Mantel-Haenszel statistics. The value of *type* can be one of the following:

- MODRIDIT
- RANK

- RIDIT
- TABLE

See the section "Scores" on page 2314 for descriptions of these score types.

If you do not specify the SCORES= option, PROC FREQ uses SCORES=TABLE by default. For character variables, the row and column TABLE scores are the row and column numbers. That is, the TABLE score is 1 for row 1, 2 for row 2, and so on. For numeric variables, the row and column TABLE scores equal the variable values. See the section "Scores" on page 2314 for details. Using MODRIDIT, RANK, or RIDIT scores yields nonparametric analyses.

You can use the SCOROUT option to display the row and column scores.

SCOROUT

displays the row and column scores that PROC FREQ uses to compute score-based tests and statistics. You can specify the score type with the SCORES= option. See the section "Scores" on page 2314 for details.

The scores are computed and displayed only when PROC FREQ computes statistics for two-way tables. You can use ODS to store the scores in an output data set. See the section "ODS Table Names" on page 2382 for more information.

SPARSE

reports all possible combinations of the variable values for an n-way table when $n > 1$, even if a combination does not occur in the data. The SPARSE option applies only to crosstabulation tables displayed in LIST format and to the OUT= output data set. If you do not use the LIST or OUT= option, the SPARSE option has no effect.

When you specify the SPARSE and LIST options, PROC FREQ displays all combinations of variable values in the table listing, including those with a frequency count of zero. By default, without the SPARSE option, PROC FREQ does not display zero-frequency levels in LIST output. When you use the SPARSE and OUT= options, PROC FREQ includes empty crosstabulation table cells in the output data set. By default, PROC FREQ does not include zero-frequency table cells in the output data set.

See the section "Missing Values" on page 2310 for more information.

TESTF=(*values*)

specifies the null hypothesis frequencies for a one-way chi-square goodness-of-fit test, which you request with the CHISQ option. See the section "Chi-Square Test for One-Way Tables" on page 2316 for details.

You can separate the TESTF= *values* with blanks or commas. The number of *values* must equal the number of variable levels in the one-way table. The sum of the *values* must equal the total frequency for the one-way table. List the *values* in the order in which the corresponding variable levels appear in the output. If you omit the CHISQ option, the TESTF= option invokes CHISQ.

TESTP=(*values*)

specifies the null hypothesis proportions for a one-way chi-square goodness-of-fit test, which you request with the CHISQ option. See the section "Chi-Square Test for One-Way Tables" on page 2316 for details.

You can separate the TESTP= *values* with blanks or commas. The number of *values* must equal the number of variable levels in the one-way table. List the *values* in the order in which the corresponding variable levels appear in the output. You can specify *values* in probability form as numbers between 0 and 1, where the proportions sum to 1. Or you can specify *values* in percentage form as numbers between 0 and 100, where the percentages sum to 100. If you omit the CHISQ option, the TESTP= option invokes CHISQ.

TOTPCT

displays the percentage of the total multiway table frequency in crosstabulation tables for n-way tables, where $n > 2$. By default, PROC FREQ displays the percentage of the individual two-way table frequency but does not display the percentage of the total frequency for multiway crosstabulation tables. See the section "Multiway Tables" on page 2376 for more information.

The percentage of total multiway table frequency is displayed by default when you specify the LIST option. It is also provided by default in the PERCENT variable in the OUT= output data set.

TREND

requests the Cochran-Armitage test for trend. The table must be $2 \times C$ or $R \times 2$ to compute the trend test. See the section "Cochran-Armitage Test for Trend" on page 2349 for details.

TEST Statement

TEST *options* **;**

The TEST statement requests asymptotic tests for measures of association and measures of agreement. You must use a TABLES statement with the TEST statement.

options

specify the statistics for which to provide asymptotic tests. Table 36.13 lists the available statistics, which include measures of association and agreement. The option names are identical to those in the TABLES and OUTPUT statements. You can request all tests for groups of statistics by using group options MEASURES or AGREE. Or you can request tests individually by using the options shown in Table 36.13.

For each measure of association or agreement that you specify, PROC FREQ provides an asymptotic test that the measure equals zero. PROC FREQ displays the asymptotic standard error under the null hypothesis, the test statistic, and the p-values. Additionally, PROC FREQ reports the confidence limits for the measure. The ALPHA= option in the TABLES statement determines the confidence level, which by default equals 0.05 and provides 95% confidence limits. See the sections "Asymptotic Tests" on page 2321 and "Confidence Limits" on page 2320 for details. Also see the section "Statistical Computations" on page 2314 for information about individual measures.

You can request exact tests for selected measures of association and agreement by using the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

If you use only one TABLES statement, you do not need to specify the same options in both the TABLES and TEST statements; when you specify an option in the TEST statement, PROC FREQ automatically invokes the corresponding TABLES statement option. However, when you use the TEST

statement with multiple TABLES statements, you must specify options in the TABLES statements to request the desired statistics. PROC FREQ then provides asymptotic tests for those statistics that you also specify in the TEST statement.

Table 36.13 TEST Statement Options

Option	Asymptotic Tests	Required TABLES Statement Option
AGREE	simple and weighted kappa coefficients	AGREE
GAMMA	gamma	ALL or MEASURES
KAPPA	simple kappa coefficient	AGREE
KENTB	Kendall's tau-b	ALL or MEASURES
MEASURES	gamma, Kendall's tau-b, Stuart's tau-c, Somers' $D(C\|R)$, Somers' $D(R\|C)$, Pearson and Spearman correlations	ALL or MEASURES
PCORR	Pearson correlation coefficient	ALL or MEASURES
SCORR	Spearman correlation coefficient	ALL or MEASURES
SMDCR	Somers' $D(C\|R)$	ALL or MEASURES
SMDRC	Somers' $D(R\|C)$	ALL or MEASURES
STUTC	Stuart's tau-c	ALL or MEASURES
WTKAP	weighted kappa coefficient	AGREE

WEIGHT Statement

WEIGHT *variable* < / *option* > **;**

The WEIGHT statement names a numeric variable that provides a weight for each observation in the input data set. The WEIGHT statement is most commonly used to input cell count data. See the section "Inputting Frequency Counts" on page 2308 for more information. If you use a WEIGHT statement, PROC FREQ assumes that an observation represents n observations, where n is the value of *variable*. The value of the WEIGHT variable is not required to be an integer.

If the value of the WEIGHT variable is missing, PROC FREQ does not use that observation in the analysis. If the value of the WEIGHT variable is zero, PROC FREQ ignores the observation unless you specify the ZEROS option, which includes observations with zero weights. If you do not specify a WEIGHT statement, PROC FREQ assigns a weight of one to each observation. The sum of the WEIGHT variable values represents the total number of observations.

If any value of the WEIGHT variable is negative, PROC FREQ displays the frequencies computed from the weighted values but does not compute percentages and statistics. If you create an output data set by using the OUT= option in the TABLES statement, PROC FREQ assigns missing values to the PERCENT variable. PROC FREQ also assigns missing values to the variables that the OUTEXPECT and OUTPCT options provide. If any value of the WEIGHT variable is negative, you cannot create an output data set by using the OUTPUT statement because statistics are not computed when there are negative weights.

You can specify the following *option* in the WEIGHT statement:

ZEROS

includes observations with zero weight values. By default, PROC FREQ ignores observations with zero weights.

If you specify the ZEROS option, frequency and and crosstabulation tables display any levels corresponding to observations with zero weights. Without the ZEROS option, PROC FREQ does not process observations with zero weights, and so does not display levels that contain only observations with zero weights.

With the ZEROS option, PROC FREQ includes levels with zero weights in the chi-square goodness-of-fit test for one-way tables. Also, PROC FREQ includes any levels with zero weights in binomial computations for one-way tables. This makes it possible to compute binomial tests and estimates when the specified level contains no observations with positive weights.

For two-way tables, the ZEROS option enables computation of kappa statistics when there are levels that contain no observations with positive weight. For more information, see the section "Tables with Zero Rows and Columns" on page 2356.

Note that even with the ZEROS option, PROC FREQ does not compute the CHISQ or MEASURES statistics for two-way tables when the table has a zero row or zero column because most of these statistics are undefined in this case.

The ZEROS option invokes the SPARSE option in the TABLES statement, which includes table cells with a zero frequency count in the LIST output and in the OUT= data set. By default, without the SPARSE option, PROC FREQ does not include zero frequency cells in the LIST output or in the OUT= data set. If you specify the ZEROS option in the WEIGHT statement but do not want the SPARSE option, you can specify the NOSPARSE option in the TABLES statement.

Details: FREQ Procedure

Inputting Frequency Counts

PROC FREQ can use either raw data or cell count data to produce frequency and crosstabulation tables. *Raw data*, also known as case-record data, report the data as one record for each subject or sample member. *Cell count data* report the data as a table, listing all possible combinations of data values along with the frequency counts. This way of presenting data often appears in published results.

The following DATA step statements store raw data in a SAS data set:

```
data Raw;
   input Subject $ R C @@;
   datalines;
01 1 1  02 1 1  03 1 1  04 1 1  05 1 1
06 1 2  07 1 2  08 1 2  09 2 1  10 2 1
11 2 1  12 2 1  13 2 2  14 2 2  14 2 2
;
```

You can store the same data as cell counts by using the following DATA step statements:

```
data CellCounts;
   input R C Count @@;
   datalines;
1 1 5   1 2 3
2 1 4   2 2 3
;
```

The variable R contains the values for the rows, and the variable C contains the values for the columns. The variable Count contains the cell count for each row and column combination.

Both the Raw data set and the CellCounts data set produce identical frequency counts, two-way tables, and statistics. When using the CellCounts data set, you must include a WEIGHT statement to specify that the variable Count contains cell counts. For example, the following PROC FREQ statements create a two-way crosstabulation table by using the CellCounts data set:

```
proc freq data=CellCounts;
   tables R*C;
   weight Count;
run;
```

Grouping with Formats

PROC FREQ groups a variable's values according to its formatted values. If you assign a format to a variable with a FORMAT statement, PROC FREQ formats the variable values before dividing observations into the levels of a frequency or crosstabulation table.

For example, suppose that variable X has the values 1.1, 1.4, 1.7, 2.1, and 2.3. Each of these values appears as a level in the frequency table. If you decide to round each value to a single digit, include the following statement in the PROC FREQ step:

```
format X 1.;
```

Now the table lists the frequency count for formatted level 1 as two and for formatted level 2 as three.

PROC FREQ treats formatted character variables in the same way. The formatted values are used to group the observations into the levels of a frequency table or crosstabulation table. PROC FREQ uses the entire value of a character format to classify an observation.

You can also use the FORMAT statement to assign formats that were created with the FORMAT procedure to the variables. User-written formats determine the number of levels for a variable and provide labels for a table. If you use the same data with different formats, then you can produce frequency counts and statistics for different classifications of the variable values.

When you use PROC FORMAT to create a user-written format that combines missing and nonmissing values into one category, PROC FREQ treats the entire category of formatted values as missing. For example, a

questionnaire codes 1 as yes, 2 as no, and 8 as a no answer. The following PROC FORMAT statements create a user-written format:

```
proc format;
   value Questfmt 1   ='Yes'
                  2   ='No'
                  8,. ='Missing';
run;
```

When you use a FORMAT statement to assign Questfmt. to a variable, the variable's frequency table no longer includes a frequency count for the response of 8. You must use the MISSING or MISSPRINT option in the TABLES statement to list the frequency for no answer. The frequency count for this level includes observations with either a value of 8 or a missing value (.).

The frequency or crosstabulation table lists the values of both character and numeric variables in ascending order based on internal (unformatted) variable values unless you change the order with the ORDER= option. To list the values in ascending order by formatted values, use ORDER=FORMATTED in the PROC FREQ statement.

For more information about the FORMAT statement, see *SAS Language Reference: Concepts*.

Missing Values

When the value of the WEIGHT variable is missing, PROC FREQ does not include that observation in the analysis.

PROC FREQ treats missing BY variable values like any other BY variable value. The missing values form a separate BY group.

If an observation has a missing value for a variable in a TABLES request, by default PROC FREQ does not include that observation in the frequency or crosstabulation table. Also by default, PROC FREQ does not include observations with missing values in the computation of percentages and statistics. The procedure displays the number of missing observations below each table.

PROC FREQ also reports the number of missing values in output data sets. The TABLES statement OUT= data set includes an observation that contains the missing value frequency. The NMISS option in the OUTPUT statement provides an output data set variable that contains the missing value frequency.

The following options change the way in which PROC FREQ handles missing values of TABLES variables:

MISSPRINT displays missing value frequencies in frequency or crosstabulation tables but does not include them in computations of percentages or statistics.

MISSING treats missing values as a valid nonmissing level for all TABLES variables. Displays missing levels in frequency and crosstabulation tables and includes them in computations of percentages and statistics.

This example shows the three ways that PROC FREQ can handle missing values of TABLES variables. The following DATA step statements create a data set with a missing value for the variable A:

```
data one;
   input A Freq;
   datalines;
1 2
2 2
. 2
;
```

The following PROC FREQ statements request a one-way frequency table for the variable A. The first request does not specify a missing value option. The second request specifies the MISSPRINT option in the TABLES statement. The third request specifies the MISSING option in the TABLES statement.

```
proc freq data=one;
   tables A;
   weight Freq;
   title 'Default';
run;
proc freq data=one;
   tables A / missprint;
   weight Freq;
   title 'MISSPRINT Option';
run;
proc freq data=one;
   tables A / missing;
   weight Freq;
   title 'MISSING Option';
run;
```

Figure 36.12 displays the frequency tables produced by this example. The first table shows PROC FREQ's default behavior for handling missing values. The observation with a missing value of the TABLES variable A is not included in the table, and the frequency of missing values is displayed below the table. The second table, for which the MISSPRINT option is specified, displays the missing observation but does not include its frequency when computing the total frequency and percentages. The third table shows that PROC FREQ treats the missing level as a valid nonmissing level when the MISSING option is specified. The table displays the missing level, and PROC FREQ includes this level when computing frequencies and percentages.

Figure 36.12 Missing Values in Frequency Tables

Default

The FREQ Procedure

A	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	2	50.00	2	50.00
2	2	50.00	4	100.00

Frequency Missing = 2

MISSPRINT Option

The FREQ Procedure

A	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	2	.	.	.
1	2	50.00	2	50.00
2	2	50.00	4	100.00

Frequency Missing = 2

MISSING Option

The FREQ Procedure

A	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	2	33.33	2	33.33
1	2	33.33	4	66.67
2	2	33.33	6	100.00

When a combination of variable values for a two-way table is missing, PROC FREQ assigns zero to the frequency count for the table cell. By default, PROC FREQ does not display missing combinations in LIST format. Also, PROC FREQ does not include missing combinations in the OUT= output data set by default. To include missing combinations, you can specify the SPARSE option with the LIST or OUT= option in the TABLES statement.

In-Database Computation

The FREQ procedure can use in-database computation to construct frequency and crosstabulation tables when the DATA= input data set is stored as a table in a supported database management system (DBMS). Supported databases include Teradata, DB2 under UNIX, and Oracle. In-database computation can provide the advantages of faster processing and reduced data transfer between the database and SAS software. For information about in-database computation, see the section "In-Database Procedures" in *SAS/ACCESS 9.2 for Relational Databases: Reference*.

PROC FREQ performs in-database computation by using SQL implicit pass-through. The procedure generates SQL queries that are based on the tables that you request in the TABLES statement. The database executes these SQL queries to construct initial summary tables, which are then transmitted to PROC FREQ. The procedure uses this summary information to perform the remaining analyses and tasks in the usual way (out of the database). So instead of transferring the entire data set over the network between the database and SAS software, the in-database method transfers only the summary tables. This can substantially reduce processing time when the dimensions of the summary tables (in terms of rows and columns) are much smaller than the dimensions of the entire database table (in terms of individual observations). Additionally, in-database summarization uses efficient parallel processing, which can also provide performance advantages.

In-database computation is controlled by the SQLGENERATION option, which you can specify in either a LIBNAME statement or an OPTIONS statement. See the section "In-Database Procedures" in *SAS/ACCESS 9.2 for Relational Databases: Reference* for details about the SQLGENERATION option and other options that affect in-database computation. By default, PROC FREQ uses in-database computation when possible. There are no FREQ procedure options that control in-database computation.

PROC FREQ uses formatted values to group observations into the levels of frequency and crosstabulation tables. See the section "Grouping with Formats" on page 2309 for more information. If formats are available in the database, then in-database summarization uses the formats. If formats are not available in the database, then in-database summarization is based on the raw data values, and PROC FREQ performs the final, formatted classification (out of the database). For more information, see the section "Deploying and Using SAS Formats in Teradata" in *SAS/ACCESS 9.2 for Relational Databases: Reference*.

The order of observations is not inherently defined for DBMS tables. The following options relate to the order of observations and therefore should not be specified for PROC FREQ in-database computation:

- If you specify the FIRSTOBS= or OBS= data set option, PROC FREQ does not perform in-database computation.
- If you specify the NOTSORTED option in the BY statement, PROC FREQ in-database computation ignores it and uses the default ASCENDING order for BY variables.
- If you specify the ORDER=DATA option for input data in a DBMS table, PROC FREQ computation might produce different results for separate runs of the same analysis. In addition to determining the order of variable levels in crosstabulation table displays, the ORDER= option can also affect the values of many of the test statistics and measures that PROC FREQ computes.

Statistical Computations

Definitions and Notation

A two-way table represents the crosstabulation of row variable X and column variable Y. Let the table row values or levels be denoted by X_i, $i = 1, 2, \ldots, R$, and the column values by Y_j, $j = 1, 2, \ldots, C$. Let n_{ij} denote the frequency of the table cell in the ith row and jth column and define the following notation:

$$n_{i\cdot} = \sum_j n_{ij} \quad \text{(row totals)}$$

$$n_{\cdot j} = \sum_i n_{ij} \quad \text{(column totals)}$$

$$n = \sum_i \sum_j n_{ij} \quad \text{(overall total)}$$

$$p_{ij} = n_{ij}/n \quad \text{(cell percentages)}$$

$$p_{i\cdot} = n_{i\cdot}/n \quad \text{(row percentages of total)}$$

$$p_{\cdot j} = n_{\cdot j}/n \quad \text{(column percentages of total)}$$

$$R_i = \text{score for row } i$$

$$C_j = \text{score for column } j$$

$$\bar{R} = \sum_i n_{i\cdot} R_i / n \quad \text{(average row score)}$$

$$\bar{C} = \sum_j n_{\cdot j} C_j / n \quad \text{(average column score)}$$

$$A_{ij} = \sum_{k>i} \sum_{l>j} n_{kl} + \sum_{k<i} \sum_{l<j} n_{kl}$$

$$D_{ij} = \sum_{k>i} \sum_{l<j} n_{kl} + \sum_{k<i} \sum_{l>j} n_{kl}$$

$$P = \sum_i \sum_j n_{ij} A_{ij} \quad \text{(twice the number of concordances)}$$

$$Q = \sum_i \sum_j n_{ij} D_{ij} \quad \text{(twice the number of discordances)}$$

Scores

PROC FREQ uses scores of the variable values to compute the Mantel-Haenszel chi-square, Pearson correlation, Cochran-Armitage test for trend, weighted kappa coefficient, and Cochran-Mantel-Haenszel statistics. The SCORES= option in the TABLES statement specifies the score type that PROC FREQ uses. The available score types are TABLE, RANK, RIDIT, and MODRIDIT scores. The default score type is TABLE. Using MODRIDIT, RANK, or RIDIT scores yields nonparametric analyses.

For numeric variables, table scores are the values of the row and column levels. If the row or column variable is formatted, then the table score is the internal numeric value corresponding to that level. If two or more numeric values are classified into the same formatted level, then the internal numeric value for that level is the smallest of these values. For character variables, table scores are defined as the row numbers and column numbers (that is, 1 for the first row, 2 for the second row, and so on).

Rank scores, which you request with the SCORES=RANK option, are defined as

$$R1_i = \sum_{k<i} n_{k\cdot} + (n_{i\cdot} + 1)/2 \qquad i = 1, 2, \ldots, R$$
$$C1_j = \sum_{l<j} n_{\cdot l} + (n_{\cdot j} + 1)/2 \qquad j = 1, 2, \ldots, C$$

where $R1_i$ is the rank score of row i, and $C1_j$ is the rank score of column j. Note that rank scores yield midranks for tied values.

Ridit scores, which you request with the SCORES=RIDIT option, are defined as rank scores standardized by the sample size (Bross 1958, Mack and Skillings 1980). Ridit scores are derived from the rank scores as

$$R2_i = R1_i/n \qquad i = 1, 2, \ldots, R$$
$$C2_j = C1_j/n \qquad j = 1, 2, \ldots, C$$

Modified ridit scores (SCORES=MODRIDIT) represent the expected values of the order statistics of the uniform distribution on (0,1) (van Elteren 1960, Lehmann 1975). Modified ridit scores are derived from rank scores as

$$R3_i = R1_i/(n+1) \qquad i = 1, 2, \ldots, R$$
$$C3_j = C1_j/(n+1) \qquad j = 1, 2, \ldots, C$$

Chi-Square Tests and Statistics

The CHISQ option provides chi-square tests of homogeneity or independence and measures of association based on the chi-square statistic. When you specify the CHISQ option in the TABLES statement, PROC FREQ computes the following chi-square tests for each two-way table: the Pearson chi-square, likelihood-ratio chi-square, and Mantel-Haenszel chi-square. PROC FREQ provides the following measures of association based on the Pearson chi-square statistic: the phi coefficient, contingency coefficient, and Cramer's V. For 2×2 tables, the CHISQ option also provides Fisher's exact test and the continuity-adjusted chi-square. You can request Fisher's exact test for general $R \times C$ tables by specifying the FISHER option in the TABLES or EXACT statement.

For one-way frequency tables, the CHISQ option provides a chi-square goodness-of-fit test. The other chi-square tests and statistics described in this section are computed only for two-way tables.

All of the two-way test statistics described in this section test the null hypothesis of no association between the row variable and the column variable. When the sample size n is large, these test statistics have an asymptotic chi-square distribution when the null hypothesis is true. When the sample size is not large, exact tests might be useful. PROC FREQ provides exact tests for the Pearson chi-square, the likelihood-ratio chi-square, and the Mantel-Haenszel chi-square (in addition to Fisher's exact test). PROC FREQ also provides an exact chi-square goodness-of-fit test for one-way tables. You can request these exact tests by specifying

the corresponding options in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Note that the Mantel-Haenszel chi-square statistic is appropriate only when both variables lie on an ordinal scale. The other chi-square tests and statistics in this section are appropriate for either nominal or ordinal variables. The following sections give the formulas that PROC FREQ uses to compute the chi-square tests and statistics. See Agresti (2007), Stokes, Davis, and Koch (2000), and the other references cited for each statistic for more information.

Chi-Square Test for One-Way Tables

For one-way frequency tables, the CHISQ option in the TABLES statement provides a chi-square goodness-of-fit test. Let C denote the number of classes, or levels, in the one-way table. Let f_i denote the frequency of class i (or the number of observations in class i) for $i = 1, 2, \ldots, C$. Then PROC FREQ computes the one-way chi-square statistic as

$$Q_P = \sum_{i=1}^{C} \frac{(f_i - e_i)^2}{e_i}$$

where e_i is the expected frequency for class i under the null hypothesis.

In the test for equal proportions, which is the default for the CHISQ option, the null hypothesis specifies equal proportions of the total sample size for each class. Under this null hypothesis, the expected frequency for each class equals the total sample size divided by the number of classes,

$$e_i = n/C \quad \text{for } i = 1, 2, \ldots, C$$

In the test for specified frequencies, which PROC FREQ computes when you input null hypothesis frequencies by using the TESTF= option, the expected frequencies are the TESTF= values that you specify. In the test for specified proportions, which PROC FREQ computes when you input null hypothesis proportions by using the TESTP= option, the expected frequencies are determined from the specified TESTP= proportions p_i as

$$e_i = p_i \times n \quad \text{for } i = 1, 2, \ldots, C$$

Under the null hypothesis (of equal proportions, specified frequencies, or specified proportions), Q_P has an asymptotic chi-square distribution with $C - 1$ degrees of freedom.

In addition to the asymptotic test, you can request an exact one-way chi-square test by specifying the CHISQ option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Pearson Chi-Square Test for Two-Way Tables

The Pearson chi-square for two-way tables involves the differences between the observed and expected frequencies, where the expected frequencies are computed under the null hypothesis of independence. The Pearson chi-square statistic is computed as

$$Q_P = \sum_i \sum_j \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

where n_{ij} is the observed frequency in table cell (i, j) and e_{ij} is the expected frequency for table cell (i, j). The expected frequency is computed under the null hypothesis that the row and column variables are independent,

$$e_{ij} = \frac{n_{i\cdot}\, n_{\cdot j}}{n}$$

When the row and column variables are independent, Q_P has an asymptotic chi-square distribution with $(R-1)(C-1)$ degrees of freedom. For large values of Q_P, this test rejects the null hypothesis in favor of the alternative hypothesis of general association.

In addition to the asymptotic test, you can request an exact Pearson chi-square test by specifying the PCHI or CHISQ option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

For 2×2 tables, the Pearson chi-square is also appropriate for testing the equality of two binomial proportions. For $R \times 2$ and $2 \times C$ tables, the Pearson chi-square tests the homogeneity of proportions. See Fienberg (1980) for details.

Likelihood-Ratio Chi-Square Test

The likelihood-ratio chi-square involves the ratios between the observed and expected frequencies. The likelihood-ratio chi-square statistic is computed as

$$G^2 = 2 \sum_i \sum_j n_{ij} \ln\left(\frac{n_{ij}}{e_{ij}}\right)$$

where n_{ij} is the observed frequency in table cell (i, j) and e_{ij} is the expected frequency for table cell (i, j).

When the row and column variables are independent, G^2 has an asymptotic chi-square distribution with $(R-1)(C-1)$ degrees of freedom.

In addition to the asymptotic test, you can request an exact likelihood-ratio chi-square test by specifying the LRCHI or CHISQ option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Continuity-Adjusted Chi-Square Test

The continuity-adjusted chi-square for 2×2 tables is similar to the Pearson chi-square, but it is adjusted for the continuity of the chi-square distribution. The continuity-adjusted chi-square is most useful for small sample sizes. The use of the continuity adjustment is somewhat controversial; this chi-square test is more conservative (and more like Fisher's exact test) when the sample size is small. As the sample size increases, the continuity-adjusted chi-square becomes more like the Pearson chi-square.

The continuity-adjusted chi-square statistic is computed as

$$Q_C = \sum_i \sum_j \frac{\left(\max(0, |n_{ij} - e_{ij}| - 0.5)\right)^2}{e_{ij}}$$

Under the null hypothesis of independence, Q_C has an asymptotic chi-square distribution with $(R-1)(C-1)$ degrees of freedom.

Mantel-Haenszel Chi-Square Test

The Mantel-Haenszel chi-square statistic tests the alternative hypothesis that there is a linear association between the row variable and the column variable. Both variables must lie on an ordinal scale. The Mantel-Haenszel chi-square statistic is computed as

$$Q_{MH} = (n-1)r^2$$

where r^2 is the Pearson correlation between the row variable and the column variable. For a description of the Pearson correlation, see the "Pearson Correlation Coefficient" on page 2324. The Pearson correlation and thus the Mantel-Haenszel chi-square statistic use the scores that you specify in the SCORES= option in the TABLES statement. See Mantel and Haenszel (1959) and Landis, Heyman, and Koch (1978) for more information.

Under the null hypothesis of no association, Q_{MH} has an asymptotic chi-square distribution with one degree of freedom.

In addition to the asymptotic test, you can request an exact Mantel-Haenszel chi-square test by specifying the MHCHI or CHISQ option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Fisher's Exact Test

Fisher's exact test is another test of association between the row and column variables. This test assumes that the row and column totals are fixed, and then uses the hypergeometric distribution to compute probabilities of possible tables conditional on the observed row and column totals. Fisher's exact test does not depend on any large-sample distribution assumptions, and so it is appropriate even for small sample sizes and for sparse tables.

2 × 2 Tables For 2 × 2 tables, PROC FREQ gives the following information for Fisher's exact test: table probability, two-sided p-value, left-sided p-value, and right-sided p-value. The table probability equals the hypergeometric probability of the observed table, and is in fact the value of the test statistic for Fisher's exact test.

Where p is the hypergeometric probability of a specific table with the observed row and column totals, Fisher's exact p-values are computed by summing probabilities p over defined sets of tables,

$$PROB = \sum_{A} p$$

The two-sided p-value is the sum of all possible table probabilties (conditional on the observed row and column totals) that are less than or equal to the observed table probability. For the two-sided p-value, the set A includes all possible tables with hypergeometric probabilities less than or equal to the probability of the observed table. A small two-sided p-value supports the alternative hypothesis of association between the row and column variables.

For 2 × 2 tables, one-sided p-values for Fisher's exact test are defined in terms of the frequency of the cell in the first row and first column of the table, the (1,1) cell. Denoting the observed (1,1) cell frequency by n_{11}, the left-sided p-value for Fisher's exact test is the probability that the (1,1) cell frequency is less than or equal to n_{11}. For the left-sided p-value, the set A includes those tables with a (1,1) cell frequency less

than or equal to n_{11}. A small left-sided p-value supports the alternative hypothesis that the probability of an observation being in the first cell is actually less than expected under the null hypothesis of independent row and column variables.

Similarly, for a right-sided alternative hypothesis, A is the set of tables where the frequency of the (1,1) cell is greater than or equal to that in the observed table. A small right-sided p-value supports the alternative that the probability of the first cell is actually greater than that expected under the null hypothesis.

Because the (1,1) cell frequency completely determines the 2×2 table when the marginal row and column sums are fixed, these one-sided alternatives can be stated equivalently in terms of other cell probabilities or ratios of cell probabilities. The left-sided alternative is equivalent to an odds ratio less than 1, where the odds ratio equals $(n_{11}n_{22}/n_{12}n_{21})$. Additionally, the left-sided alternative is equivalent to the column 1 risk for row 1 being less than the column 1 risk for row 2, $p_{1|1} < p_{1|2}$. Similarly, the right-sided alternative is equivalent to the column 1 risk for row 1 being greater than the column 1 risk for row 2, $p_{1|1} > p_{1|2}$. See Agresti (2007) for details.

R × C Tables Fisher's exact test was extended to general $R \times C$ tables by Freeman and Halton (1951), and this test is also known as the Freeman-Halton test. For $R \times C$ tables, the two-sided p-value definition is the same as for 2×2 tables. The set A contains all tables with p less than or equal to the probability of the observed table. A small p-value supports the alternative hypothesis of association between the row and column variables. For $R \times C$ tables, Fisher's exact test is inherently two-sided. The alternative hypothesis is defined only in terms of general, and not linear, association. Therefore, Fisher's exact test does not have right-sided or left-sided p-values for general $R \times C$ tables.

For $R \times C$ tables, PROC FREQ computes Fisher's exact test by using the network algorithm of Mehta and Patel (1983), which provides a faster and more efficient solution than direct enumeration. See the section "Exact Statistics" on page 2366 for more details.

Phi Coefficient

The phi coefficient is a measure of association derived from the Pearson chi-square. The range of the phi coefficient is $-1 \leq \phi \leq 1$ for 2×2 tables. For tables larger than 2×2, the range is $0 \leq \phi \leq \min(\sqrt{R-1}, \sqrt{C-1})$ (Liebetrau 1983). The phi coefficient is computed as

$$\phi = (n_{11}n_{22} - n_{12}n_{21}) \,/\, \sqrt{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}} \quad \text{for } 2 \times 2 \text{ tables}$$

$$\phi = \sqrt{Q_P/n} \quad \text{otherwise}$$

See Fleiss, Levin, and Paik (2003, pp. 98–99) for more information.

Contingency Coefficient

The contingency coefficient is a measure of association derived from the Pearson chi-square. The range of the contingency coefficient is $0 \leq P \leq \sqrt{(m-1)/m}$, where $m = \min(R, C)$ (Liebetrau 1983). The contingency coefficient is computed as

$$P = \sqrt{Q_P \,/\, (Q_P + n)}$$

See Kendall and Stuart (1979, pp. 587–588) for more information.

Cramer's V

Cramer's V is a measure of association derived from the Pearson chi-square. It is designed so that the attainable upper bound is always 1. The range of Cramer's V is $-1 \leq V \leq 1$ for 2×2 tables; for tables larger than 2×2, the range is $0 \leq V \leq 1$. Cramer's V is computed as

$$V = \phi \quad \text{for } 2 \times 2 \text{tables}$$

$$V = \sqrt{\frac{Q_P/n}{\min(R-1, C-1)}} \quad \text{otherwise}$$

See Kendall and Stuart (1979, p. 588) for more information.

Measures of Association

When you specify the MEASURES option in the TABLES statement, PROC FREQ computes several statistics that describe the association between the row and column variables of the contingency table. The following are measures of ordinal association that consider whether the column variable Y tends to increase as the row variable X increases: gamma, Kendall's tau-b, Stuart's tau-c, and Somers' D. These measures are appropriate for ordinal variables, and they classify pairs of observations as *concordant* or *discordant*. A pair is concordant if the observation with the larger value of X also has the larger value of Y. A pair is discordant if the observation with the larger value of X has the smaller value of Y. See Agresti (2007) and the other references cited for the individual measures of association.

The Pearson correlation coefficient and the Spearman rank correlation coefficient are also appropriate for ordinal variables. The Pearson correlation describes the strength of the linear association between the row and column variables, and it is computed by using the row and column scores specified by the SCORES= option in the TABLES statement. The Spearman correlation is computed with rank scores. The polychoric correlation (requested by the PLCORR option) also requires ordinal variables and assumes that the variables have an underlying bivariate normal distribution. The following measures of association do not require ordinal variables and are appropriate for nominal variables: lambda asymmetric, lambda symmetric, and the uncertainty coefficients.

PROC FREQ computes estimates of the measures according to the formulas given in the following sections. For each measure, PROC FREQ computes an asymptotic standard error (*ASE*), which is the square root of the asymptotic variance denoted by var in the following sections.

Confidence Limits

If you specify the CL option in the TABLES statement, PROC FREQ computes asymptotic confidence limits for all MEASURES statistics. The confidence coefficient is determined according to the value of the ALPHA= option, which, by default, equals 0.05 and produces 95% confidence limits.

The confidence limits are computed as

$$est \ \pm \ (\ z_{\alpha/2} \times ASE \ )$$

where est is the estimate of the measure, $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution, and *ASE* is the asymptotic standard error of the estimate.

Asymptotic Tests

For each measure that you specify in the TEST statement, PROC FREQ computes an asymptotic test of the null hypothesis that the measure equals zero. Asymptotic tests are available for the following measures of association: gamma, Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$, Somers' $D(R|C)$, the Pearson correlation coefficient, and the Spearman rank correlation coefficient. To compute an asymptotic test, PROC FREQ uses a standardized test statistic z, which has an asymptotic standard normal distribution under the null hypothesis. The test statistic is computed as

$$z = est \;/\; \sqrt{\mathrm{var}_0(est)}$$

where est is the estimate of the measure and $\mathrm{var}_0(est)$ is the variance of the estimate under the null hypothesis. Formulas for $\mathrm{var}_0(est)$ for the individual measures of association are given in the following sections.

Note that the ratio of est to $\sqrt{\mathrm{var}_0(est)}$ is the same for the following measures: gamma, Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$, and Somers' $D(R|C)$. Therefore, the tests for these measures are identical. For example, the p-values for the test of H_0: gamma $= 0$ equal the p-values for the test of H_0: tau-$b = 0$.

PROC FREQ computes one-sided and two-sided p-values for each of these tests. When the test statistic z is greater than its null hypothesis expected value of zero, PROC FREQ displays the right-sided p-value, which is the probability of a larger value of the statistic occurring under the null hypothesis. A small right-sided p-value supports the alternative hypothesis that the true value of the measure is greater than zero. When the test statistic is less than or equal to zero, PROC FREQ displays the left-sided p-value, which is the probability of a smaller value of the statistic occurring under the null hypothesis. A small left-sided p-value supports the alternative hypothesis that the true value of the measure is less than zero. The one-sided p-value P_1 can be expressed as

$$P_1 = \begin{cases} \mathrm{Prob}(Z > z) & \text{if } z > 0 \\ \mathrm{Prob}(Z < z) & \text{if } z \leq 0 \end{cases}$$

where Z has a standard normal distribution. The two-sided p-value P_2 is computed as

$$P_2 = \mathrm{Prob}(|Z| > |z|)$$

Exact Tests

Exact tests are available for the following measures of association: Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$ and $(R|C)$, the Pearson correlation coefficient, and the Spearman rank correlation coefficient. If you request an exact test for a measure of association in the EXACT statement, PROC FREQ computes the exact test of the hypothesis that the measure equals zero. See the section "Exact Statistics" on page 2366 for details.

Gamma

The gamma (Γ) statistic is based only on the number of concordant and discordant pairs of observations. It ignores tied pairs (that is, pairs of observations that have equal values of X or equal values of Y). Gamma is appropriate only when both variables lie on an ordinal scale. The range of gamma is $-1 \leq \Gamma \leq 1$. If the row and column variables are independent, then gamma tends to be close to zero. Gamma is estimated by

$$G = (P - Q) \;/\; (P + Q)$$

and the asymptotic variance is

$$\mathrm{var}(G) = \frac{16}{(P+Q)^4} \sum_i \sum_j n_{ij}(QA_{ij} - PD_{ij})^2$$

For 2×2 tables, gamma is equivalent to Yule's Q. See Goodman and Kruskal (1979) and Agresti (2002) for more information.

The variance under the null hypothesis that gamma equals zero is computed as

$$\mathrm{var}_0(G) = \frac{4}{(P+Q)^2} \left(\sum_i \sum_j n_{ij}(A_{ij} - D_{ij})^2 - (P-Q)^2/n \right)$$

See Brown and Benedetti (1977) for details.

Kendall's Tau-b

Kendall's tau-b (τ_b) is similar to gamma except that tau-b uses a correction for ties. Tau-b is appropriate only when both variables lie on an ordinal scale. The range of tau-b is $-1 \leq \tau_b \leq 1$. Kendall's tau-b is estimated by

$$t_b = (P-Q) \,/\, \sqrt{w_r w_c}$$

and the asymptotic variance is

$$\mathrm{var}(t_b) = \frac{1}{w^4} \left(\sum_i \sum_j n_{ij}(2wd_{ij} + t_b v_{ij})^2 - n^3 t_b^2 (w_r + w_c)^2 \right)$$

where

$$\begin{aligned} w &= \sqrt{w_r w_c} \\ w_r &= n^2 - \sum_i n_{i\cdot}^2 \\ w_c &= n^2 - \sum_j n_{\cdot j}^2 \\ d_{ij} &= A_{ij} - D_{ij} \\ v_{ij} &= n_{i\cdot} w_c + n_{\cdot j} w_r \end{aligned}$$

See Kendall (1955) for more information.

The variance under the null hypothesis that tau-b equals zero is computed as

$$\mathrm{var}_0(t_b) = \frac{4}{w_r w_c} \left(\sum_i \sum_j n_{ij}(A_{ij} - D_{ij})^2 - (P-Q)^2/n \right)$$

See Brown and Benedetti (1977) for details.

PROC FREQ also provides an exact test for the Kendall's tau-b. You can request this test by specifying the KENTB option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Stuart's Tau-c

Stuart's tau-c (τ_c) makes an adjustment for table size in addition to a correction for ties. Tau-c is appropriate only when both variables lie on an ordinal scale. The range of tau-c is $-1 \leq \tau_c \leq 1$. Stuart's tau-c is estimated by

$$t_c = m(P - Q) \,/\, n^2(m-1)$$

and the asymptotic variance is

$$\text{var}(t_c) = \frac{4m^2}{(m-1)^2 n^4} \left(\sum_i \sum_j n_{ij} d_{ij}^2 - (P-Q)^2/n \right)$$

where $m = \min(R, C)$ and $d_{ij} = A_{ij} - D_{ij}$. The variance under the null hypothesis that tau-c equals zero is the same as the asymptotic variance var,

$$\text{var}_0(t_c) = \text{var}(t_c)$$

See Brown and Benedetti (1977) for details.

PROC FREQ also provides an exact test for the Stuart's tau-c. You can request this test by specifying the STUTC option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Somers' D

Somers' $D(C|R)$ and Somers' $D(R|C)$ are asymmetric modifications of tau-b. $C|R$ indicates that the row variable X is regarded as the independent variable and the column variable Y is regarded as dependent. Similarly, $R|C$ indicates that the column variable Y is regarded as the independent variable and the row variable X is regarded as dependent. Somers' D differs from tau-b in that it uses a correction only for pairs that are tied on the independent variable. Somers' D is appropriate only when both variables lie on an ordinal scale. The range of Somers' D is $-1 \leq D \leq 1$. Somers' $D(C|R)$ is computed as

$$D(C|R) = (P - Q) \,/\, w_r$$

and its asymptotic variance is

$$\text{var}(D(C|R)) = \frac{4}{w_r^4} \sum_i \sum_j n_{ij} \left(w_r d_{ij} - (P-Q)(n - n_{i.}) \right)^2$$

where $d_{ij} = A_{ij} - D_{ij}$ and

$$w_r = n^2 - \sum_i n_{i.}^2$$

See Somers (1962), Goodman and Kruskal (1979), and Liebetrau (1983) for more information.

The variance under the null hypothesis that $D(C|R)$ equals zero is computed as

$$\mathrm{var}_0(D(C|R)) = \frac{4}{w_r^2}\left(\sum_i\sum_j n_{ij}(A_{ij} - D_{ij})^2 - (P-Q)^2/n\right)$$

See Brown and Benedetti (1977) for details.

Formulas for Somers' $D(R|C)$ are obtained by interchanging the indices.

PROC FREQ also provides exact tests for Somers' $D(C|R)$ and $(R|C)$. You can request these tests by specifying the SMDCR and SMDCR options in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Pearson Correlation Coefficient

The Pearson correlation coefficient (ρ) is computed by using the scores specified in the SCORES= option. This measure is appropriate only when both variables lie on an ordinal scale. The range of the Pearson correlation is $-1 \le \rho \le 1$. The Pearson correlation coefficient is estimated by

$$r = v/w = ss_{rc}/\sqrt{ss_r ss_c}$$

and its asymptotic variance is

$$\mathrm{var}(r) = \frac{1}{w^4}\sum_i\sum_j n_{ij}\left(w(R_i - \bar{R})(C_j - \bar{C}) - \frac{b_{ij}v}{2w}\right)^2$$

where R_i and C_j are the row and column scores and

$$\begin{aligned} ss_r &= \sum_i\sum_j n_{ij}(R_i - \bar{R})^2 \\ ss_c &= \sum_i\sum_j n_{ij}(C_j - \bar{C})^2 \\ ss_{rc} &= \sum_i\sum_j n_{ij}(R_i - \bar{R})(C_j - \bar{C}) \end{aligned}$$

$$\begin{aligned} b_{ij} &= (R_i - \bar{R})^2 ss_c + (C_j - \bar{C})^2 ss_r \\ v &= ss_{rc} \\ w &= \sqrt{ss_r ss_c} \end{aligned}$$

See Snedecor and Cochran (1989) for more information.

The SCORES= option in the TABLES statement determines the type of row and column scores used to compute the Pearson correlation (and other score-based statistics). The default is SCORES=TABLE. See the section "Scores" on page 2314 for details about the available score types and how they are computed.

The variance under the null hypothesis that the correlation equals zero is computed as

$$\mathrm{var}_0(r) = \left(\sum_i \sum_j n_{ij} (R_i - \bar{R})^2 (C_j - \bar{C})^2 - ss_{rc}^2 / n \right) \; / \; ss_r ss_c$$

Note that this expression for the variance is derived for multinomial sampling in a contingency table framework, and it differs from the form obtained under the assumption that both variables are continuous and normally distributed. See Brown and Benedetti (1977) for details.

PROC FREQ also provides an exact test for the Pearson correlation coefficient. You can request this test by specifying the PCORR option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Spearman Rank Correlation Coefficient

The Spearman correlation coefficient (ρ_s) is computed by using rank scores, which are defined in the section "Scores" on page 2314. This measure is appropriate only when both variables lie on an ordinal scale. The range of the Spearman correlation is $-1 \leq \rho_s \leq 1$. The Spearman correlation coefficient is estimated by

$$r_s = v \; / \; w$$

and its asymptotic variance is

$$\mathrm{var}(r_s) = \frac{1}{n^2 w^4} \sum_i \sum_j n_{ij} (z_{ij} - \bar{z})^2$$

where $R1_i$ and $C1_j$ are the row and column rank scores and

$$\begin{aligned} v &= \sum_i \sum_j n_{ij} R(i) C(j) \\ w &= \frac{1}{12} \sqrt{FG} \\ F &= n^3 - \sum_i n_{i\cdot}^3 \\ G &= n^3 - \sum_j n_{\cdot j}^3 \end{aligned}$$

$$
\begin{aligned}
R(i) &= R1_i - n/2 \\
C(j) &= C1_j - n/2 \\
\bar{z} &= \frac{1}{n}\sum_i \sum_j n_{ij} z_{ij} \\
z_{ij} &= w v_{ij} - v w_{ij} \\
v_{ij} &= n\left(R(i)C(j) + \frac{1}{2}\sum_l n_{il} C(l) + \frac{1}{2}\sum_k n_{kj} R(k) + \right. \\
&\qquad \left. \sum_l \sum_{k>i} n_{kl} C(l) + \sum_k \sum_{l>j} n_{kl} R(k) \right) \\
w_{ij} &= \frac{-n}{96w}\left(F n_{\cdot j}^2 + G n_{i\cdot}^2 \right)
\end{aligned}
$$

See Snedecor and Cochran (1989) for more information.

The variance under the null hypothesis that the correlation equals zero is computed as

$$\mathrm{var}_0(r_s) = \frac{1}{n^2 w^2} \sum_i \sum_j n_{ij} (v_{ij} - \bar{v})^2$$

where

$$\bar{v} = \sum_i \sum_j n_{ij} v_{ij} / n$$

Note that the asymptotic variance is derived for multinomial sampling in a contingency table framework, and it differs from the form obtained under the assumption that both variables are continuous and normally distributed. See Brown and Benedetti (1977) for details.

PROC FREQ also provides an exact test for the Spearman correlation coefficient. You can request this test by specifying the SCORR option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Polychoric Correlation

When you specify the PLCORR option in the TABLES statement, PROC FREQ computes the polychoric correlation. This measure of association is based on the assumption that the ordered, categorical variables of the frequency table have an underlying bivariate normal distribution. For 2×2 tables, the polychoric correlation is also known as the tetrachoric correlation. See Drasgow (1986) for an overview of polychoric correlation. The polychoric correlation coefficient is the maximum likelihood estimate of the product-moment correlation between the normal variables, estimating thresholds from the observed table frequencies. The range of the polychoric correlation is from –1 to 1. Olsson (1979) gives the likelihood equations and an asymptotic covariance matrix for the estimates.

To estimate the polychoric correlation, PROC FREQ iteratively solves the likelihood equations by a Newton-Raphson algorithm that uses the Pearson correlation coefficient as the initial approximation. Iteration stops when the convergence measure falls below the convergence criterion or when the maximum number of iterations is reached, whichever occurs first. The CONVERGE= option sets the convergence criterion, and the default value is 0.0001. The MAXITER= option sets the maximum number of iterations, and the default value is 20.

Lambda (Asymmetric)

Asymmetric lambda, $\lambda(C|R)$, is interpreted as the probable improvement in predicting the column variable Y given knowledge of the row variable X. The range of asymmetric lambda is $0 \leq \lambda(C|R) \leq 1$. Asymmetric lambda $(C|R)$ is computed as

$$\lambda(C|R) = \frac{\sum_i r_i - r}{n - r}$$

and its asymptotic variance is

$$\text{var}(\lambda(C|R)) = \frac{n - \sum_i r_i}{(n-r)^3}\left(\sum_i r_i + r - 2\sum_i (r_i \mid l_i = l)\right)$$

where

$$\begin{aligned} r_i &= \max_j(n_{ij}) \\ r &= \max_j(n_{\cdot j}) \\ c_j &= \max_i(n_{ij}) \\ c &= \max_i(n_{i\cdot}) \end{aligned}$$

The values of l_i and l are determined as follows. Denote by l_i the unique value of j such that $r_i = n_{ij}$, and let l be the unique value of j such that $r = n_{\cdot j}$. Because of the uniqueness assumptions, ties in the frequencies or in the marginal totals must be broken in an arbitrary but consistent manner. In case of ties, l is defined as the smallest value of j such that $r = n_{\cdot j}$.

For those columns containing a cell (i, j) for which $n_{ij} = r_i = c_j$, cs_j records the row in which c_j is assumed to occur. Initially cs_j is set equal to –1 for all j. Beginning with $i = 1$, if there is at least one value j such that $n_{ij} = r_i = c_j$, and if $cs_j = -1$, then l_i is defined to be the smallest such value of j, and cs_j is set equal to i. Otherwise, if $n_{il} = r_i$, then l_i is defined to be equal to l. If neither condition is true, then l_i is taken to be the smallest value of j such that $n_{ij} = r_i$.

The formulas for lambda asymmetric $(R|C)$ can be obtained by interchanging the indices.

See Goodman and Kruskal (1979) for more information.

Lambda (Symmetric)

The nondirectional lambda is the average of the two asymmetric lambdas, $\lambda(C|R)$ and $\lambda(R|C)$. Its range is $0 \leq \lambda \leq 1$. Lambda symmetric is computed as

$$\lambda = \frac{\sum_i r_i + \sum_j c_j - r - c}{2n - r - c} = \frac{w - v}{w}$$

and its asymptotic variance is computed as

$$\text{var}(\lambda) = \frac{1}{w^4}\Big(wvy - 2w^2\big(n - \sum_i \sum_j (n_{ij} \mid j = l_i, i = k_j)\big) - 2v^2(n - n_{kl})\Big)$$

where

$$\begin{aligned}
r_i &= \max_j(n_{ij}) \\
r &= \max_j(n_{.j}) \\
c_j &= \max_i(n_{ij}) \\
c &= \max_i(n_{i.}) \\
w &= 2n - r - c \\
v &= 2n - \sum_i r_i - \sum_j c_j \\
x &= \sum_i (r_i \mid l_i = l) + \sum_j (c_j \mid k_j = k) + r_k + c_l \\
y &= 8n - w - v - 2x
\end{aligned}$$

The definitions of l_i and l are given in the previous section. The values k_j and k are defined in a similar way for lambda asymmetric $(R|C)$.

See Goodman and Kruskal (1979) for more information.

Uncertainty Coefficients (Asymmetric)

The uncertainty coefficient $U(C|R)$ measures the proportion of uncertainty (entropy) in the column variable Y that is explained by the row variable X. Its range is $0 \leq U(C|R) \leq 1$. The uncertainty coefficient is computed as

$$U(C|R) = (H(X) + H(Y) - H(XY)) \;/\; H(Y) = v/w$$

and its asymptotic variance is

$$\text{var}(U(C|R)) = \frac{1}{n^2 w^4} \sum_i \sum_j n_{ij}\Big(H(Y)\ln\Big(\frac{n_{ij}}{n_{i.}}\Big) + (H(X) - H(XY))\ln\Big(\frac{n_{.j}}{n}\Big)\Big)^2$$

where

$$\begin{aligned} v &= H(X) + H(Y) - H(XY) \\ w &= H(Y) \\ H(X) &= -\sum_i \left(\frac{n_{i\cdot}}{n}\right) \ln\left(\frac{n_{i\cdot}}{n}\right) \\ H(Y) &= -\sum_j \left(\frac{n_{\cdot j}}{n}\right) \ln\left(\frac{n_{\cdot j}}{n}\right) \\ H(XY) &= -\sum_i \sum_j \left(\frac{n_{ij}}{n}\right) \ln\left(\frac{n_{ij}}{n}\right) \end{aligned}$$

The formulas for the uncertainty coefficient $U(R|C)$ can be obtained by interchanging the indices.

See Theil (1972, pp. 115–120) and Goodman and Kruskal (1979) for more information.

Uncertainty Coefficient (Symmetric)

The uncertainty coefficient U is the symmetric version of the two asymmetric uncertainty coefficients. Its range is $0 \leq U \leq 1$. The uncertainty coefficient is computed as

$$U = 2\left(H(X) + H(Y) - H(XY)\right) \;/\; \left(H(X) + H(Y)\right)$$

and its asymptotic variance is

$$\text{var}(U) = 4 \sum_i \sum_j \frac{n_{ij}\left(H(XY)\ln\left(\frac{n_{i\cdot}n_{\cdot j}}{n^2}\right) - \left(H(X)+H(Y)\right)\ln\left(\frac{n_{ij}}{n}\right)\right)^2}{n^2\left(H(X)+H(Y)\right)^4}$$

where $H(X)$, $H(Y)$, and $H(XY)$ are defined in the previous section. See Goodman and Kruskal (1979) for more information.

Binomial Proportion

If you specify the BINOMIAL option in the TABLES statement, PROC FREQ computes the binomial proportion for one-way tables. By default, this is the proportion of observations in the first variable level that appears in the output. (You can use the LEVEL= option to specify a different level for the proportion.) The binomial proportion is computed as

$$\hat{p} = n_1 \;/\; n$$

where n_1 is the frequency of the first (or designated) level and n is the total frequency of the one-way table. The standard error of the binomial proportion is computed as

$$se(\hat{p}) = \sqrt{\hat{p}\left(1-\hat{p}\right) \;/\; n}$$

Binomial Confidence Limits

By default, PROC FREQ provides asymptotic and exact (Clopper-Pearson) confidence limits for the binomial proportion. If you do not specify any confidence limit requests with *binomial-options*, PROC FREQ computes the standard Wald asymptotic confidence limits. You can also request Agresti-Coull, Jeffreys, and Wilson (score) confidence limits for the binomial proportion. See Brown, Cai, and DasGupta (2001), Agresti and Coull (1998), and Newcombe (1998) for details about these binomial confidence limits, including comparisons of their performance.

Wald Confidence Limits The standard Wald asymptotic confidence limits are based on the normal approximation to the binomial distribution. PROC FREQ computes the Wald confidence limits for the binomial proportion as

$$\hat{p} \pm (z_{\alpha/2} \times \mathrm{se}(\hat{p}))$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution. The confidence level α is determined by the ALPHA= option, which, by default, equals 0.05 and produces 95% confidence limits.

If you specify the CORRECT *binomial-option* or the BINOMIALC option, PROC FREQ includes a continuity correction of $1/2n$ in the Wald asymptotic confidence limits. The purpose of this correction is to adjust for the difference between the normal approximation and the binomial distribution, which is a discrete distribution. See Fleiss, Levin, and Paik (2003) for more information. With the continuity correction, the asymptotic confidence limits for the binomial proportion are computed as

$$\hat{p} \pm (z_{\alpha/2} \times \mathrm{se}(\hat{p}) + (1/2n))$$

Agresti-Coull Confidence Limits If you specify the AGRESTICOULL *binomial-option*, PROC FREQ computes Agresti-Coull confidence limits for the binomial proportion as

$$\tilde{p} \pm (z_{\alpha/2} \times \sqrt{\tilde{p}\,(1-\tilde{p}) \,/\, \tilde{n}})$$

where

$$\begin{aligned} \tilde{n}_1 &= n_1 + (z_{\alpha/2})/2 \\ \tilde{n} &= n + z^2_{\alpha/2} \\ \tilde{p} &= \tilde{n}_1 \,/\, \tilde{n} \end{aligned}$$

The Agresti-Coull confidence interval has the same basic form as the standard Wald interval but uses $\tilde{p}$ in place of $\hat{p}$. For $\alpha = 0.05$, the value of $z_{\alpha/2}$ is close to 2, and this interval is the "add 2 successes and 2 failures" adjusted Wald interval in Agresti and Coull (1998).

Jeffreys Confidence Limits If you specify the JEFFREYS *binomial-option*, PROC FREQ computes the Jeffreys confidence limits for the binomial proportion as

$$\big(\beta(\alpha/2,\ n_1 + 1/2,\ n - n_1 + 1/2),\ \ \beta(1-\alpha/2,\ n_1 + 1/2,\ n - n_1 + 1/2) \big)$$

where $\beta(\alpha, b, c)$ is the αth percentile of the beta distribution with shape parameters b and c. The lower confidence limit is set to 0 when $n_1 = 0$, and the upper confidence limit is set to 1 when $n_1 = n$. This is an

equal-tailed interval based on the noninformative Jeffreys prior for a binomial proportion. See Brown, Cai, and DasGupta (2001) for details. See Berger (1985) for information about using beta priors for inference on the binomial proportion.

Wilson (Score) Confidence Limits If you specify the WILSON *binomial-option*, PROC FREQ computes Wilson confidence limits for the binomial proportion. These are also known as score confidence limits and are attributed to Wilson (1927). The confidence limits are based on inverting the normal test that uses the null proportion in the variance (the score test). Wilson confidence limits are the roots of

$$|p - \hat{p}| = z_{\alpha/2}\sqrt{p(1-p)/n}$$

and are computed as

$$\left(\hat{p} \;+\; z^2_{\alpha/2}/2n \;\pm\; z_{\alpha/2}\sqrt{\left(\hat{p}(1-\hat{p}) + z^2_{\alpha/2}/4n\right)/n}\right) \;/\; \left(1 + z^2_{\alpha/2}/n\right)$$

The Wilson interval has been shown to have better performance than the Wald interval and the exact (Clopper-Pearson) interval. See Agresti and Coull (1998), Brown, Cai, and DasGupta (2001), and Newcombe (1998) for more information.

Exact (Clopper-Pearson) Confidence Limits Exact (Clopper-Pearson) confidence limits for the binomial proportion are constructed by inverting the equal-tailed test based on the binomial distribution. This method is attributed to Clopper and Pearson (1934). The exact confidence limits p_L and p_U satisfy the following equations, for $n_1 = 1, 2, \ldots n-1$:

$$\sum_{x=n_1}^{n} \binom{n}{x} p_L^{\,x}(1-p_L)^{\,n-x} \;=\; \alpha/2$$

$$\sum_{x=0}^{n_1} \binom{n}{x} p_U^{\,x}(1-p_U)^{\,n-x} \;=\; \alpha/2$$

The lower confidence limit equals 0 when $n_1 = 0$, and the upper confidence limit equals 1 when $n_1 = n$.

PROC FREQ computes the exact (Clopper-Pearson) confidence limits by using the F distribution as

$$p_L \;=\; \left(1 + \frac{n - n_1 + 1}{n_1\, F(1-\alpha/2,\; 2n_1,\; 2(n-n_1+1))}\right)^{-1}$$

$$p_U \;=\; \left(1 + \frac{n - n_1}{(n_1+1)\, F(\alpha/2,\; 2(n_1+1),\; 2(n-n_1))}\right)^{-1}$$

where $F(\alpha, b, c)$ is the αth percentile of the F distribution with b and c degrees of freedom. See Leemis and Trivedi (1996) for a derivation of this expression. Also see Collett (1991) for more information about exact binomial confidence limits.

Because this is a discrete problem, the confidence coefficient (or coverage probability) of the exact (Clopper-Pearson) interval is not exactly $(1-\alpha)$ but is at least $(1-\alpha)$. Thus, this confidence interval is conservative. Unless the sample size is large, the actual coverage probability can be much larger than the target value. See Agresti and Coull (1998), Brown, Cai, and DasGupta (2001), and Leemis and Trivedi (1996) for more information about the performance of these confidence limits.

Binomial Tests

The BINOMIAL option provides an asymptotic equality test for the binomial proportion by default. You can also specify *binomial-options* to request tests of noninferiority, superiority, and equivalence for the binomial proportion. If you specify the BINOMIAL option in the EXACT statement, PROC FREQ also computes exact p-values for the tests that you request with the *binomial-options*.

Equality Test PROC FREQ computes an asymptotic test of the hypothesis that the binomial proportion equals p_0, where you can specify the value of p_0 with the P= *binomial-option*. If you do not specify a null value with P=, PROC FREQ uses $p_0 = 0.5$ by default. The binomial test statistic is computed as

$$z = (\hat{p} - p_0)/se$$

By default, the standard error is based on the null hypothesis proportion as

$$se = \sqrt{p_0(1-p_0)/n}$$

If you specify the VAR=SAMPLE *binomial-option*, the standard error is computed from the sample proportion as

$$se = \sqrt{\hat{p}(1-\hat{p})/n}$$

If you specify the CORRECT *binomial-option* or the BINOMIALC option, PROC FREQ includes a continuity correction in the asymptotic test statistic, towards adjusting for the difference between the normal approximation and the discrete binomial distribution. See Fleiss, Levin, and Paik (2003) for details. The continuity correction of $(1/2n)$ is subtracted from the numerator of the test statistic if $(\hat{p} - p_0)$ is positive; otherwise, the continuity correction is added to the numerator.

PROC FREQ computes one-sided and two-sided p-values for this test. When the test statistic z is greater than zero (its expected value under the null hypothesis), PROC FREQ computes the right-sided p-value, which is the probability of a larger value of the statistic occurring under the null hypothesis. A small right-sided p-value supports the alternative hypothesis that the true value of the proportion is greater than p_0. When the test statistic is less than or equal to zero, PROC FREQ computes the left-sided p-value, which is the probability of a smaller value of the statistic occurring under the null hypothesis. A small left-sided p-value supports the alternative hypothesis that the true value of the proportion is less than p_0. The one-sided p-value P_1 can be expressed as

$$P_1 = \begin{cases} \text{Prob}(Z > z) & \text{if } z > 0 \\ \text{Prob}(Z < z) & \text{if } z \leq 0 \end{cases}$$

where Z has a standard normal distribution. The two-sided p-value is computed as $P_2 = 2 \times P_1$.

If you specify the BINOMIAL option in the EXACT statement, PROC FREQ also computes an exact test of the null hypothesis H_0: $p = p_0$. To compute the exact test, PROC FREQ uses the binomial probability function,

$$\text{Prob}(X = x \mid p_0) = \binom{n}{x} {p_0}^x \, (1 - p_0)^{(n-x)} \quad \text{for} \quad x = 0, 1, 2, \ldots, n$$

where the variable X has a binomial distribution with parameters n and p_0. To compute the left-sided p-value, $\text{Prob}(X \leq n_1)$, PROC FREQ sums the binomial probabilities over x from zero to n_1. To compute the right-sided p-value, $\text{Prob}(X \geq n_1)$, PROC FREQ sums the binomial probabilities over x from n_1 to n. The exact one-sided p-value is the minimum of the left-sided and right-sided p-values,

$$P_1 = \min \left(\text{Prob}(X \leq n_1 \mid p_0), \ \text{Prob}(X \geq n_1 \mid p_0) \right)$$

and the exact two-sided p-value is computed as $P_2 = 2 \times P_1$.

Noninferiority Test If you specify the NONINF *binomial-option*, PROC FREQ provides a noninferiority test for the binomial proportion. The null hypothesis for the noninferiority test is

$$H_0: p - p_0 \leq -\delta$$

versus the alternative

$$H_a: p - p_0 > -\delta$$

where δ is the noninferiority margin and p_0 is the null proportion. Rejection of the null hypothesis indicates that the binomial proportion is not inferior to the null value. See Chow, Shao, and Wang (2003) for more information.

You can specify the value of δ with the MARGIN= *binomial-option*, and you can specify p_0 with the P= *binomial-option*. By default, $\delta = 0.2$ and $p_0 = 0.5$.

PROC FREQ provides an asymptotic Wald test for noninferiority. The test statistic is computed as

$$z = (\hat{p} - p_0^*) \ / \ se$$

where p_0^* is the noninferiority limit,

$$p_0^* = p_0 - \delta$$

By default, the standard error is computed from the sample proportion as

$$se = \sqrt{\hat{p}(1 - \hat{p})/n}$$

If you specify the VAR=NULL *binomial-option*, the standard error is based on the noninferiority limit (determined by the null proportion and the margin) as

$$se = \sqrt{p_0^*(1 - p_0^*)/n}$$

If you specify the CORRECT *binomial-option* or the BINOMIALC option, PROC FREQ includes a continuity correction in the asymptotic test statistic z. The continuity correction of $(1/2n)$ is subtracted from the numerator of the test statistic if $(\hat{p} - p_0^*)$ is positive; otherwise, the continuity correction is added to the numerator.

The p-value for the noninferiority test is

$$P_z = \text{Prob}(Z > z)$$

where Z has a standard normal distribution.

As part of the noninferiority analysis, PROC FREQ provides asymptotic Wald confidence limits for the binomial proportion. These confidence limits are computed as described in the section "Wald Confidence Limits" on page 2330 but use the same standard error (VAR=NULL or VAR=SAMPLE) as the noninferiority test statistic z. The confidence coefficient is $100(1 - 2\alpha)\%$ (Schuirmann 1999). By default, if you do not specify the ALPHA= option, the noninferiority confidence limits are 90% confidence limits. You can compare the confidence limits to the noninferiority limit, $p_0^* = p_0 - \delta$.

If you specify the BINOMIAL option in the EXACT statement, PROC FREQ provides an exact noninferiority test for the binomial proportion. The exact p-value is computed by using the binomial probability function with parameters p_0^* and n,

$$P_x = \sum_{k=n_1}^{k=n} \binom{n}{k} (p_0^*)^{\,k}\, (1 - p_0^*)^{\,(n-k)}$$

See Chow, Shao, Wang (2003, p. 116) for details. If you request exact binomial statistics, PROC FREQ also includes exact (Clopper-Pearson) confidence limits for the binomial proportion in the equivalence analysis display. See the section "Exact (Clopper-Pearson) Confidence Limits" on page 2331 for details.

Superiority Test If you specify the SUP *binomial-option*, PROC FREQ provides a superiority test for the binomial proportion. The null hypothesis for the superiority test is

$$H_0\colon p - p_0 \leq \delta$$

versus the alternative

$$H_a\colon p - p_0 > \delta$$

where δ is the superiority margin and p_0 is the null proportion. Rejection of the null hypothesis indicates that the binomial proportion is superior to the null value. You can specify the value of δ with the MARGIN= *binomial-option*, and you can specify the value of p_0 with the P= *binomial-option*. By default, $\delta = 0.2$ and $p_0 = 0.5$.

The superiority analysis is identical to the noninferiority analysis but uses a positive value of the margin δ in the null hypothesis. The superiority limit equals $p_0 + \delta$. The superiority computations follow those in the section "Noninferiority Test" on page 2333 but replace $-\delta$ with δ. See Chow, Shao, and Wang (2003) for more information.

Equivalence Test If you specify the EQUIV *binomial-option*, PROC FREQ provides an equivalence test for the binomial proportion. The null hypothesis for the equivalence test is

$$H_0: p - p_0 \leq \delta_L \quad \text{or} \quad p - p_0 \geq \delta_U$$

versus the alternative

$$H_a: \delta_L < p - p_0 < \delta_U$$

where δ_L is the lower margin, δ_U is the upper margin, and p_0 is the null proportion. Rejection of the null hypothesis indicates that the binomial proportion is equivalent to the null value. See Chow, Shao, and Wang (2003) for more information.

You can specify the value of the margins δ_L and δ_U with the MARGIN= *binomial-option*. If you do not specify MARGIN=, PROC FREQ uses lower and upper margins of –0.2 and 0.2 by default. If you specify a single margin value δ, PROC FREQ uses lower and upper margins of $-\delta$ and δ. You can specify the null proportion p_0 with the P= *binomial-option*. By default, $p_0 = 0.5$.

PROC FREQ computes two one-sided tests (TOST) for equivalence analysis (Schuirmann 1987). The TOST approach includes a right-sided test for the lower margin and a left-sided test for the upper margin. The overall p-value is taken to be the larger of the two p-values from the lower and upper tests.

For the lower margin, the asymptotic Wald test statistic is computed as

$$z_L = (\hat{p} - p_L^*) \,/\, se$$

where the lower equivalence limit is

$$p_L^* = p_0 + \delta_L$$

By default, the standard error is computed from the sample proportion as

$$se = \sqrt{\hat{p}(1 - \hat{p})/n}$$

If you specify the VAR=NULL *binomial-option*, the standard error is based on the lower equivalence limit (determined by the null proportion and the lower margin) as

$$se = \sqrt{p_L^*(1 - p_L^*)/n}$$

If you specify the CORRECT *binomial-option* or the BINOMIALC option, PROC FREQ includes a continuity correction in the asymptotic test statistic z_L. The continuity correction of $(1/2n)$ is subtracted from the numerator of the test statistic $(\hat{p} - p_L^*)$ if the numerator is positive; otherwise, the continuity correction is added to the numerator.

The p-value for the lower margin test is

$$P_{z,L} = \text{Prob}(Z > z_L)$$

The asymptotic test for the upper margin is computed similarly. The Wald test statistic is

$$z_U = (\hat{p} - p_U^*) \,/\, se$$

where the upper equivalence limit is

$$p_U^* = p_0 + \delta_U$$

By default, the standard error is computed from the sample proportion. If you specify the VAR=NULL *binomial-option*, the standard error is based on the upper equivalence limit as

$$se = \sqrt{p_U^*(1 - p_U^*)/n}$$

If you specify the CORRECT *binomial-option* or the BINOMIALC option, PROC FREQ includes a continuity correction of $(1/2n)$ in the asymptotic test statistic z_U.

The p-value for the upper margin test is

$$P_{z,U} = \text{Prob}(Z < z_U)$$

Based on the two one-sided tests (TOST), the overall p-value for the test of equivalence equals the larger p-value from the lower and upper margin tests, which can be expressed as

$$P_z = \max(P_{z,L}\ ,\ P_{z,U})$$

As part of the equivalence analysis, PROC FREQ provides asymptotic Wald confidence limits for the binomial proportion. These confidence limits are computed as described in the section "Wald Confidence Limits" on page 2330, but use the same standard error (VAR=NULL or VAR=SAMPLE) as the equivalence test statistics and have a confidence coefficient of $100(1 - 2\alpha)\%$ (Schuirmann 1999). By default, if you do not specify the ALPHA= option, the equivalence confidence limits are 90% limits. If you specify VAR=NULL, separate standard errors are computed for the lower and upper margin tests, each based on the null proportion and the corresponding (lower or upper) margin. The confidence limits are computed by using the maximum of these two standard errors. You can compare the confidence limits to the equivalence limits, $(p_0 + \delta_L,\ p_0 + \delta_U)$.

If you specify the BINOMIAL option in the EXACT statement, PROC FREQ also provides an exact equivalence test by using two one-sided exact tests (TOST). The procedure computes lower and upper margin exact tests by using the binomial probability function as described in the section "Noninferiority Test" on page 2333. The overall exact p-value for the equivalence test is taken to be the larger p-value from the lower and upper margin exact tests. If you request exact statistics, PROC FREQ also includes exact (Clopper-Pearson) confidence limits in the equivalence analysis display. The confidence coefficient is $100(1 - 2\alpha)\%$ (Schuirmann 1999). See the section "Exact (Clopper-Pearson) Confidence Limits" on page 2331 for details.

Risks and Risk Differences

The RISKDIFF option in the TABLES statement provides estimates of risks (binomial proportions) and risk differences for 2×2 tables. This analysis might be appropriate when comparing the proportion of some characteristic for two groups, where row 1 and row 2 correspond to the two groups, and the columns correspond to two possible characteristics or outcomes. For example, the row variable might be a treatment or dose, and the column variable might be the response. See Collett (1991), Fleiss, Levin, and Paik (2003), and Stokes, Davis, and Koch (2000) for more information.

Let the frequencies of the 2×2 table be represented as follows.

	Column 1	Column 2	Total
Row 1	n_{11}	n_{12}	$n_{1\cdot}$
Row 2	n_{21}	n_{22}	$n_{2\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	n

For column 1 and column 2, PROC FREQ provides estimates of the row 1 risk (proportion), the row 2 risk, the overall risk, and the risk difference. The risk difference is defined as the row 1 risk minus the row 2 risk. The risks are binomial proportions of their rows (row 1, row 2, or overall), and the computation of their standard errors and confidence limits follow the binomial proportion computations, which are described in the section "Binomial Proportion" on page 2329.

The column 1 risk for row 1 is the proportion of row 1 observations classified in column 1,

$$p_1 = n_{11} \; / \; n_{1\cdot}$$

This estimates the conditional probability of the column 1 response, given the first level of the row variable.

The column 1 risk for row 2 is the proportion of row 2 observations classified in column 1,

$$p_2 = n_{21} \; / \; n_{2\cdot}$$

The overall column 1 risk is the proportion of all observations classified in column 1,

$$p = n_{\cdot 1} \; / \; n$$

The column 1 risk difference compares the risks for the two rows, and it is computed as the column 1 risk for row 1 minus the column 1 risk for row 2,

$$d = p_1 - p_2$$

The risks and risk difference are defined similarly for column 2.

The standard error of the column 1 risk for row i is computed as

$$se(p_i) = \sqrt{p_i \; (1 - p_i) \; / \; n_{1\cdot}}$$

The standard error of the overall column 1 risk is computed as

$$se(p) = \sqrt{p \; (1 - p) \; / \; n}$$

If the two rows represent independent binomial samples, the standard error for the column 1 risk difference is computed as

$$se(d) = \sqrt{var(p_1) + var(p_2)}$$

The standard errors are computed in a similar manner for the column 2 risks and risk difference.

Confidence Limits

By default, the RISKDIFF option provides standard Wald asymptotic confidence limits for the risks (row 1, row 2, and overall) and the risk difference. The RISKDIFF option also provides other types of confidence limits and tests for the risk difference. See the sections "Risk Difference Confidence Limits" on page 2338 and "Risk Difference Tests" on page 2340 for details.

The risks are equivalent to binomial proportions of their corresponding rows. This section describes the Wald confidence limits for risks that are provided by the RISKDIFF option. The BINOMIAL option provides additional confidence limit types and tests for risks in the binomial proportion framework. See the sections "Binomial Confidence Limits" on page 2330 and "Binomial Tests" on page 2332 for details.

The Wald asymptotic confidence limits are based on the normal approximation to the binomial distribution. PROC FREQ computes the Wald confidence limits for the risks and risk differences as

$$est \pm (z_{\alpha/2} \times \mathrm{se}(est))$$

where est is the estimate, $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution, and $\mathrm{se}(est)$ is the standard error of the estimate. The confidence level α is determined from the value of the ALPHA= option, which, by default, equals 0.05 and produces 95% confidence limits.

If you specify the CORRECT *riskdiff-option* or the RISKDIFFC option, PROC FREQ includes continuity corrections in the Wald asymptotic confidence limits for the risks and risk differences. The purpose of a continuity correction is to adjust for the difference between the normal approximation and the binomial distribution, which is discrete. See Fleiss, Levin, and Paik (2003) for more information. With the continuity correction, the asymptotic confidence limits are computed as

$$est \pm (z_{\alpha/2} \times \mathrm{se}(est) + cc)$$

where cc is the continuity correction. For the row 1 risk, $cc = (1/2n_{1.})$; for the row 2 risk, $cc = (1/2n_{2.})$; for the overall risk, $cc = (1/2n)$; and for the risk difference, $cc = ((1/n_{1.} + 1/n_{2.})/2)$. The column 1 and column 2 risks use the same continuity corrections.

PROC FREQ also computes exact (Clopper-Pearson) confidence limits for the column 1, column 2, and overall risks. These confidence limits are constructed by inverting the equal-tailed test based on the binomial distribution. PROC FREQ uses the F distribution to compute the Clopper-Pearson confidence limits. See the section "Exact (Clopper-Pearson) Confidence Limits" on page 2331 for details.

Risk Difference Confidence Limits You can request additional confidence limits for the risk difference by specifying the CL= *riskdiff-option*. Available confidence limit types include exact unconditional, Farrington-Manning, Hauck-Anderson, Newcombe score, and Wald. Continuity-corrected versions of the Newcombe and Wald confidence limits are available. By default, the Wald confidence limits use a sample-based variance; alternatively, you can request a test-based variance and specify the null risk difference value.

The confidence coefficient for the confidence limits produced by the CL= *riskdiff-option* is $100(1-\alpha)\%$, where the value of α is determined by the ALPHA= option. By default, ALPHA=0.05, which produces 95% confidence limits. This differs from the test-based confidence limits that are provided with the equivalence, noninferiority, and superiority tests, which have a confidence coefficient of $100(1-2\alpha)\%$ (Schuirmann 1999). See the section "Risk Difference Tests" on page 2340 for details.

The section "Exact Unconditional Confidence Limits for the Risk Difference" on page 2345 describes the computation of the exact confidence limits. The confidence limits are constructed by inverting two separate

one-sided exact tests (tail method). By default, the tests are based on the unstandardized risk difference. If you specify the RISKDIFF(METHOD=FMSCORE) option, the Farrington-Manning score is used as the test statistic.

PROC FREQ computes the Newcombe confidence limits for the risk difference as described in the subsection Newcombe Score Confidence Limits in the section "Noninferiority Tests" on page 2341, except that the Newcombe confidence limits produced by the CL= *riskdiff-option* have a confidence coefficient of $100(1-\alpha)\%$.

The following sections describe the computation of the Farrington-Manning, Hauck-Anderson, and Wald confidence limits for the risk difference.

Farrington-Manning Confidence Limits The Farrington-Manning confidence limits for the risk difference are computed as

$$\hat{d} \pm (z_{\alpha/2} \times \mathrm{se}(\hat{d}))$$

where $\hat{d} = \hat{p}_1 - \hat{p}_2$, $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution, and the standard error is

$$\mathrm{se}(\hat{d}) = \sqrt{\tilde{p}_1(1-\tilde{p}_1)/n_{1\cdot} + \tilde{p}_2(1-\tilde{p}_2)/n_{2\cdot}}$$

where $\tilde{p}_1$ and $\tilde{p}_2$ are the maximum likelihood estimators of p_1 and p_2 under the null hypothesis that the risk difference equals d_0.

The subsection Farrington-Manning Test in the section "Noninferiority Tests" on page 2341 describes the computation of the maximum likelihood estimators $\tilde{p}_1$ and $\tilde{p}_2$. See Farrington and Manning (1990) for details.

This computation uses a null hypothesis value of the risk difference, which you can specify in the CL=FM(NULL=*value*) *riskdiff-option*. By default, PROC FREQ uses a null value of 0. This differs from the Farrington-Manning confidence limits that are produced in the noninferiority analysis, where the null value of the risk difference is based on the test margin (which is specified by the MARGIN= *riskdiff-option*).

Hauck-Anderson Confidence Limits The Hauck-Anderson confidence limits for the risk difference are computed as

$$\hat{d} \pm (cc + z_{\alpha/2} \times \mathrm{se}(\hat{d}))$$

where $\hat{d} = \hat{p}_1 - \hat{p}_2$ and $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution. The standard error is computed from the sample proportions as

$$\mathrm{se}(\hat{d}) = \sqrt{\hat{p}_1(1-\hat{p}_1)/(n_{1\cdot}-1) + \hat{p}_2(1-\hat{p}_2)/(n_{2\cdot}-1)}$$

The Hauck-Anderson continuity correction cc is computed as

$$cc = 1 \,/\, \left(2 \min(n_{1\cdot}, n_{2\cdot})\right)$$

See Hauck and Anderson (1986) for more information. The subsection Hauck-Anderson Test in the section "Noninferiority Tests" on page 2341 describes the corresponding noninferiority test.

Wald Confidence Limits The Wald confidence limits for the risk difference are computed as

$$\hat{d} \pm (z_{\alpha/2} \times \mathrm{se}(\hat{d}))$$

where $\hat{d} = \hat{p}_1 - \hat{p}_2$ and $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution. By default, the standard error is computed from the sample proportions as

$$\mathrm{se}(\hat{d}) = \sqrt{\hat{p}_1(1-\hat{p}_1)/n_{1\cdot} + \hat{p}_2(1-\hat{p}_2)/n_{2\cdot}}$$

If you specify the CL=WALD(NULL=*value*) *riskdiff-option*, the standard error is based on the null hypothesis that the risk difference equals $d_0 =$ *value* (Dunnett and Gent 1977). The standard error is computed as

$$\mathrm{se}(\hat{d}) = \sqrt{\tilde{p}(1-\tilde{p})/n_{2\cdot} + (\tilde{p}+d_0)(1-\tilde{p}-d_0)/n_{1\cdot}}$$

where

$$\tilde{p} = (n_{11} + n_{21} + d_0 n_{1\cdot})/n$$

If you specify the CORRECT *riskdiff-option*, the Wald confidence limits include a continuity correction cc,

$$\hat{d} \pm (cc + z_{\alpha/2} \times \mathrm{se}(\hat{d}))$$

where $cc = (1/n_{1\cdot} + 1/n_{2\cdot})/2$.

The subsection Wald Test in the section "Noninferiority Tests" on page 2341 describes the corresponding noninferiority test.

Risk Difference Tests

You can specify *riskdiff-options* to request tests of the risk (proportion) difference. You can request tests of equality, noninferiority, superiority, and equivalence for the risk difference. The test of equality is a standard Wald asymptotic test, available with or without a continuity correction. For noninferiority, superiority, and equivalence tests of the risk difference, the following test methods are provided: Wald (with and without continuity correction), Hauck-Anderson, Farrington-Manning, and Newcombe score (with and without continuity correction). You can specify the test method with the METHOD= *riskdiff-option*. By default, PROC FREQ uses METHOD=WALD.

Equality Test If you specify the EQUAL *riskdiff-option*, PROC FREQ computes a test of equality, or a test of the null hypothesis that the risk difference equals zero. For the column 1 (or 2) risk difference, this test can be expressed as $H_0: d = 0$ versus the alternative $H_a: d \neq 0$, where $d = p_1 - p_2$ denotes the column 1 (or 2) risk difference. PROC FREQ provides a Wald asymptotic test of equality. The test statistic is computed as

$$z = \hat{d}/\mathrm{se}(\hat{d})$$

By default, the standard error is computed from the sample proportions as

$$\mathrm{se}(\hat{d}) = \sqrt{\hat{p}_1(1-\hat{p}_1)/n_{1\cdot} + \hat{p}_2(1-\hat{p}_2)/n_{2\cdot}}$$

If you specify the VAR=NULL *riskdiff-option*, the standard error is based on the null hypothesis that the row 1 and row 2 risks are equal,

$$\mathrm{se}(\hat{d}) = \sqrt{\hat{p}(1-\hat{p}) \times (1/n_{1\cdot} + 1/n_{2\cdot})}$$

where $\hat{p} = n_{\cdot 1}/n$ estimates the overall column 1 risk.

If you specify the CORRECT *riskdiff-option* or the RISKDIFFC option, PROC FREQ includes a continuity correction in the test statistic. If $\hat{d} > 0$, the continuity correction is subtracted from $\hat{d}$ in the numerator of the test statistic; otherwise, the continuity correction is added to the numerator. The value of the continuity correction is $(1/n_{1\cdot} + 1/n_{2\cdot})/2$.

PROC FREQ computes one-sided and two-sided p-values for this test. When the test statistic z is greater than 0, PROC FREQ displays the right-sided p-value, which is the probability of a larger value occurring under the null hypothesis. The one-sided p-value can be expressed as

$$P_1 = \begin{cases} \text{Prob}(Z > z) & \text{if } z > 0 \\ \text{Prob}(Z < z) & \text{if } z \leq 0 \end{cases}$$

where Z has a standard normal distribution. The two-sided p-value is computed as $P_2 = 2 \times P_1$.

Noninferiority Tests If you specify the NONINF *riskdiff-option*, PROC FREQ provides a noninferiority test for the risk difference, or the difference between two proportions. The null hypothesis for the noninferiority test is

$$H_0\colon p_1 - p_2 \leq -\delta$$

versus the alternative

$$H_a\colon p_1 - p_2 > -\delta$$

where δ is the noninferiority margin. Rejection of the null hypothesis indicates that the row 1 risk is not inferior to the row 2 risk. See Chow, Shao, and Wang (2003) for more information.

You can specify the value of δ with the MARGIN= *riskdiff-option*. By default, $\delta = 0.2$. You can specify the test method with the METHOD= *riskdiff-option*. The following methods are available for the risk difference noninferiority analysis: Wald (with and without continuity correction), Hauck-Anderson, Farrington-Manning, and Newcombe score (with and without continuity correction). The Wald, Hauck-Anderson, and Farrington-Manning methods provide tests and corresponding test-based confidence limits; the Newcombe score method provides only confidence limits. If you do not specify METHOD=, PROC FREQ uses the Wald test by default.

The confidence coefficient for the test-based confidence limits is $100(1 - 2\alpha)\%$ (Schuirmann 1999). By default, if you do not specify the ALPHA= option, these are 90% confidence limits. You can compare the confidence limits to the noninferiority limit, $-\delta$.

The following sections describe the noninferiority analysis methods for the risk difference.

Wald Test If you specify the METHOD=WALD *riskdiff-option*, PROC FREQ provides an asymptotic Wald test of noninferiority for the risk difference. This is also the default method. The Wald test statistic is computed as

$$z = (\hat{d} + \delta) \;/\; \text{se}(\hat{d})$$

where $(\hat{d} = \hat{p}_1 - \hat{p}_2)$ estimates the risk difference and δ is the noninferiority margin.

By default, the standard error for the Wald test is computed from the sample proportions as

$$\text{se}(\hat{d}) = \sqrt{\hat{p}_1(1 - \hat{p}_1)/n_{1\cdot} \;+\; \hat{p}_2(1 - \hat{p}_2)/n_{2\cdot}}$$

If you specify the VAR=NULL *riskdiff-option*, the standard error is based on the null hypothesis that the risk difference equals –δ (Dunnett and Gent 1977). The standard error is computed as

$$\mathrm{se}(\hat{d}) = \sqrt{\tilde{p}(1-\tilde{p})/n_{2\cdot} \; + \; (\tilde{p}-\delta)(1-\tilde{p}+\delta)/n_{1\cdot}}$$

where

$$\tilde{p} = (n_{11} + n_{21} + \delta n_{1\cdot})/n$$

If you specify the CORRECT *riskdiff-option* or the RISKDIFFC option, a continuity correction is included in the test statistic. The continuity correction is subtracted from the numerator of the test statistic if the numerator is greater than zero; otherwise, the continuity correction is added to the numerator. The value of the continuity correction is $(1/n_{1\cdot} + 1/n_{2\cdot})/2$.

The p-value for the Wald noninferiority test is $P_z = \mathrm{Prob}(Z > z)$, where Z has a standard normal distribution.

Hauck-Anderson Test If you specify the METHOD=HA *riskdiff-option*, PROC FREQ provides the Hauck-Anderson test for noninferiority. The Hauck-Anderson test statistic is computed as

$$z = (\hat{d} + \delta \; \pm \; cc) \; / \; \mathrm{se}(\hat{d})$$

where $\hat{d} = \hat{p}_1 - \hat{p}_2$ and the standard error is computed from the sample proportions as

$$\mathrm{se}(\hat{d}) = \sqrt{\hat{p}_1(1-\hat{p}_1)/(n_{1\cdot}-1) \; + \; \hat{p}_2(1-\hat{p}_2)/(n_{2\cdot}-1)}$$

The Hauck-Anderson continuity correction cc is computed as

$$cc = 1 \; / \; \left(2 \; \min(n_{1\cdot}, \; n_{2\cdot})\right)$$

The p-value for the Hauck-Anderson noninferiority test is $P_z = \mathrm{Prob}(Z > z)$, where Z has a standard normal distribution. See Hauck and Anderson (1986) and Schuirmann (1999) for more information.

Farrington-Manning Test If you specify the METHOD=FM *riskdiff-option*, PROC FREQ provides the Farrington-Manning test of noninferiority for the risk difference. The Farrington-Manning test statistic is computed as

$$z = (\hat{d} + \delta) \; / \; \mathrm{se}(\hat{d})$$

where $\hat{d} = \hat{p}_1 - \hat{p}_2$ and

$$\mathrm{se}(\hat{d}) = \sqrt{\tilde{p}_1(1-\tilde{p}_1)/n_{1\cdot} \; + \; \tilde{p}_2(1-\tilde{p}_2)/n_{2\cdot}}$$

where $\tilde{p}_1$ and $\tilde{p}_2$ are the maximum likelihood estimators of p_1 and p_2 under the null hypothesis that the risk difference equals –δ. The p-value for the Farrington-Manning noninferiority test is then $P_z = \mathrm{Prob}(Z > z)$, where Z has a standard normal distribution.

From Farrington and Manning (1990), the solution to the maximum likelihood equation is

$$\tilde{p}_1 = 2u\cos(w) - b/3a \quad \text{and} \quad \tilde{p}_2 = \tilde{p}_1 + \delta$$

where

$$
\begin{aligned}
w &= (\pi + \cos^{-1}(v/u^3))/3 \\
v &= b^3/(3a)^3 - bc/6a^2 + d/2a \\
u &= \text{sign}(v)\sqrt{b^2/(3a)^2 - c/3a} \\
a &= 1 + \theta \\
b &= -(1 + \theta + \hat{p}_1 + \theta\hat{p}_2 - \delta(\theta + 2)) \\
c &= \delta^2 - \delta(2\hat{p}_1 + \theta + 1) + \hat{p}_1 + \theta\hat{p}_2 \\
d &= \hat{p}_1\delta(1 - \delta) \\
\theta &= n_{2\cdot}/n_{1\cdot}
\end{aligned}
$$

Newcombe Score Confidence Limits If you specify the METHOD=NEWCOMBE *riskdiff-option*, PROC FREQ provides the Newcombe hybrid score (Wilson) confidence limits for the risk difference. The confidence coefficient for the confidence limits is $100(1 - 2\alpha)\%$ (Schuirmann 1999). By default, if you do not specify the ALPHA= option, these are 90% confidence limits. You can compare the confidence limits with the noninferiority limit, $-\delta$.

The Newcombe score confidence limits for the risk difference are constructed from the Wilson score confidence limits for each of the two individual proportions. The confidence limits for the individual proportions are used in the standard error terms of the Wald confidence limits for the proportion difference. See Newcombe (1998) and Barker et al. (2001) for more information.

Wilson score confidence limits for p_1 and p_2 are the roots of

$$|p_i - \hat{p}_i| = z_\alpha\sqrt{p_i(1 - p_i)/n_{i\cdot}}$$

for $i = 1, 2$. The confidence limits are computed as

$$\left(\hat{p}_i + z_\alpha^2/2n_{i\cdot} \pm z_\alpha\sqrt{\left(\hat{p}_i(1 - \hat{p}_i) + z_\alpha^2/4n_{i\cdot}\right)/n_{i\cdot}}\right) / \left(1 + z_\alpha^2/n_{i\cdot}\right)$$

See the section "Wilson (Score) Confidence Limits" on page 2331 for details.

Denote the lower and upper Wilson score confidence limits for p_1 as L_1 and U_1, and denote the lower and upper confidence limits for p_2 as L_2 and U_2. The Newcombe score confidence limits for the proportion difference ($d = p_1 - p_2$) are computed as

$$d_L = (\hat{p}_1 - \hat{p}_2) - \sqrt{(\hat{p}_1 - L_1)^2 + (U_2 - \hat{p}_2)^2}$$

$$d_U = (\hat{p}_1 - \hat{p}_2) + \sqrt{(U_1 - \hat{p}_1)^2 + (\hat{p}_2 - L_2)^2}$$

If you specify the CORRECT *riskdiff-option*, PROC FREQ provides continuity-corrected Newcombe score confidence limits. By including a continuity correction of $1/2n_{i\cdot}$, the Wilson score confidence limits for the individual proportions are computed as the roots of

$$|p_i - \hat{p}_i| - 1/2n_{i\cdot} = z_\alpha\sqrt{p_i(1 - p_i)/n_{i\cdot}}$$

The continuity-corrected confidence limits for the individual proportions are then used to compute the proportion difference confidence limits d_L and d_U.

Superiority Test If you specify the SUP *riskdiff-option*, PROC FREQ provides a superiority test for the risk difference. The null hypothesis is

$$H_0: p_1 - p_2 \leq \delta$$

versus the alternative

$$H_a: p_1 - p_2 > \delta$$

where δ is the superiority margin. Rejection of the null hypothesis indicates that the row 1 proportion is superior to the row 2 proportion. You can specify the value of δ with the MARGIN= *riskdiff-option*. By default, $\delta = 0.2$.

The superiority analysis is identical to the noninferiority analysis but uses a positive value of the margin δ in the null hypothesis. The superiority computations follow those in the section "Noninferiority Tests" on page 2341 by replacing $-\delta$ by δ. See Chow, Shao, and Wang (2003) for more information.

Equivalence Tests If you specify the EQUIV *riskdiff-option*, PROC FREQ provides an equivalence test for the risk difference, or the difference between two proportions. The null hypothesis for the equivalence test is

$$H_0: p_1 - p_2 \leq -\delta_L \quad \text{or} \quad p_1 - p_2 \geq \delta_U$$

versus the alternative

$$H_a: \delta_L < p_1 - p_2 < \delta_U$$

where δ_L is the lower margin and δ_U is the upper margin. Rejection of the null hypothesis indicates that the two binomial proportions are equivalent. See Chow, Shao, and Wang (2003) for more information.

You can specify the value of the margins δ_L and δ_U with the MARGIN= *riskdiff-option*. If you do not specify MARGIN=, PROC FREQ uses lower and upper margins of –0.2 and 0.2 by default. If you specify a single margin value δ, PROC FREQ uses lower and upper margins of $-\delta$ and δ. You can specify the test method with the METHOD= *riskdiff-option*. The following methods are available for the risk difference equivalence analysis: Wald (with and without continuity correction), Hauck-Anderson, Farrington-Manning, and Newcombe's score (with and without continuity correction). The Wald, Hauck-Anderson, and Farrington-Manning methods provide tests and corresponding test-based confidence limits; the Newcombe score method provides only confidence limits. If you do not specify METHOD=, PROC FREQ uses the Wald test by default.

PROC FREQ computes two one-sided tests (TOST) for equivalence analysis (Schuirmann 1987). The TOST approach includes a right-sided test for the lower margin δ_L and a left-sided test for the upper margin δ_U. The overall p-value is taken to be the larger of the two p-values from the lower and upper tests.

The section "Noninferiority Tests" on page 2341 gives details about the Wald, Hauck-Anderson, Farrington-Manning and Newcombe score methods for the risk difference. The lower margin equivalence test statistic takes the same form as the noninferiority test statistic but uses the lower margin value δ_L in place of $-\delta$.

The upper margin equivalence test statistic take the same form as the noninferiority test statistic but uses the upper margin value δ_U in place of $-\delta$.

The test-based confidence limits for the risk difference are computed according to the equivalence test method that you select. If you specify METHOD=WALD with VAR=NULL, or METHOD=FM, separate standard errors are computed for the lower and upper margin tests. In this case, the test-based confidence limits are computed by using the maximum of these two standard errors. The confidence limits have a confidence coefficient of $100(1-2\alpha)\%$ (Schuirmann 1999). By default, if you do not specify the ALPHA= option, these are 90% confidence limits. You can compare the confidence limits to the equivalence limits, (δ_L, δ_U).

Exact Unconditional Confidence Limits for the Risk Difference

If you specify the RISKDIFF option in the EXACT statement, PROC FREQ provides exact unconditional confidence limits for the risk difference. PROC FREQ computes the confidence limits by inverting two separate one-sided tests (tail method), where the size of each test is at most $\alpha/2$ and the confidence coefficient is at least $(1-\alpha)$. Exact conditional methods, described in the section "Exact Statistics" on page 2366, do not apply to the risk difference due to the presence of a nuisance parameter (Agresti 1992). The unconditional approach eliminates the nuisance parameter by maximizing the p-value over all possible values of the parameter (Santner and Snell 1980).

By default, PROC FREQ uses the unstandardized risk difference as the test statistic in the confidence limit computations. If you specify the RISKDIFF(METHOD=FMSCORE) option, the procedure uses the Farrington-Manning score statistic (Chan and Zhang 1999). The score statistic is a less discrete statistic than the raw risk difference and produces less conservative confidence limits (Agresti and Min 2001). See also Santner et al. (2007) for comparisons of methods for computing exact confidence limits for the risk difference.

PROC FREQ computes the confidence limits as follows. The risk difference is defined as the difference between the row 1 and row 2 risks (proportions), $d = p_1 - p_2$, and n_1 and n_2 denote the row totals of the 2×2 table. The joint probability function for the table can be expressed in terms of the table cell frequencies, the risk difference, and the nuisance parameter p_2 as

$$f(n_{11}, n_{21}; n_1, n_2, d, p_2) = \binom{n_1}{n_{11}} (d + p_2)^{n}_{11} (1 - d - p_2)^{n_1 - n_{11}} \times \binom{n_2}{n_{21}} p_2^{n_{21}} (1 - p_2)^{n_2 - n_{21}}$$

The $100(1-\alpha/2)\%$ confidence limits for the risk difference are computed as

$$\begin{aligned} d_L &= \sup\left(d_* : P_U(d_*) > \alpha/2\right) \\ d_U &= \inf\left(d_* : P_L(d_*) > \alpha/2\right) \end{aligned}$$

where

$$\begin{aligned} P_U(d_*) &= \sup_{p_2}\Big(\sum_{A, T(a)\geq t_0} f(n_{11}, n_{21}; n_1, n_2, d_*, p_2)\Big) \\ P_L(d_*) &= \sup_{p_2}\Big(\sum_{A, T(a)\leq t_0} f(n_{11}, n_{21}; n_1, n_2, d_*, p_2)\Big) \end{aligned}$$

The set A includes all 2×2 tables with row sums equal to n_1 and n_2, and $T(a)$ denotes the value of the test statistic for table a in A. To compute $P_U(d_*)$, the sum includes probabilities of those tables for which $(T(a) \geq t_0)$, where t_0 is the value of the test statistic for the observed table. For a fixed value of d_*, $P_U(d_*)$ is taken to be the maximum sum over all possible values of p_2.

By default, PROC FREQ uses the unstandardized risk difference as the test statistic T. If you specify the RISKDIFF(METHOD=FMSCORE) option, the procedure uses the Farrington-Manning risk difference score statistic as the test statistic. The computation of the risk difference score statistic is described in the subsection Farrington-Manning Test in the section "Noninferiority Tests" on page 2341. See Farrington and Manning (1990) and Miettinen and Nurminen (1985) for more information.

Odds Ratio and Relative Risks for 2 x 2 Tables

Odds Ratio (Case-Control Studies)

The odds ratio is a useful measure of association for a variety of study designs. For a retrospective design called a *case-control study*, the odds ratio can be used to estimate the relative risk when the probability of positive response is small (Agresti 2002). In a case-control study, two independent samples are identified based on a binary (yes-no) response variable, and the conditional distribution of a binary explanatory variable is examined, within fixed levels of the response variable. See Stokes, Davis, and Koch (2000) and Agresti (2007).

The odds of a positive response (column 1) in row 1 is n_{11}/n_{12}. Similarly, the odds of a positive response in row 2 is n_{21}/n_{22}. The odds ratio is formed as the ratio of the row 1 odds to the row 2 odds. The odds ratio for a 2×2 table is defined as

$$OR = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}\, n_{22}}{n_{12}\, n_{21}}$$

The odds ratio can be any nonnegative number. When the row and column variables are independent, the true value of the odds ratio equals 1. An odds ratio greater than 1 indicates that the odds of a positive response are higher in row 1 than in row 2. Values less than 1 indicate the odds of positive response are higher in row 2. The strength of association increases with the deviation from 1.

The transformation $G = (OR - 1)/(OR + 1)$ transforms the odds ratio to the range $(-1, 1)$ with $G = 0$ when $OR = 1$; $G = -1$ when $OR = 0$; and G approaches 1 as OR approaches infinity. G is the gamma statistic, which PROC FREQ computes when you specify the MEASURES option.

The asymptotic $100(1 - \alpha)\%$ confidence limits for the odds ratio are

$$\left(OR \times \exp(-z\sqrt{v}),\ \ OR \times \exp(z\sqrt{v}) \right)$$

where

$$v = \text{var}(\ln OR) = \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}$$

and z is the $100(1-\alpha/2)$th percentile of the standard normal distribution. If any of the four cell frequencies are zero, the estimates are not computed.

Exact Confidence Limits for the Odds Ratio When you specify the OR option in the EXACT statement, PROC FREQ computes exact confidence limits for the odds ratio. Because this is a discrete problem, the confidence coefficient for the exact confidence interval is not exactly $(1 - \alpha)$ but is at least $(1 - \alpha)$. Thus, these confidence limits are conservative. See Agresti (1992) for more information.

PROC FREQ computes exact confidence limits for the odds ratio by using an algorithm based on Thomas (1971). See also Gart (1971). The following two equations are solved iteratively to determine the lower and upper confidence limits, ϕ_1 and ϕ_2:

$$\sum_{i=n_{11}}^{n_{\cdot 1}} \binom{n_{1\cdot}}{i} \binom{n_{2\cdot}}{n_{\cdot 1} - i} \phi_1^i \;/\; \sum_{i=0}^{n_{\cdot 1}} \binom{n_{1\cdot}}{i} \binom{n_{2\cdot}}{n_{\cdot 1} - i} \phi_1^i \quad = \quad \alpha/2$$

$$\sum_{i=0}^{n_{11}} \binom{n_{1\cdot}}{i} \binom{n_{2\cdot}}{n_{\cdot 1} - i} \phi_2^i \;/\; \sum_{i=0}^{n_{\cdot 1}} \binom{n_{1\cdot}}{i} \binom{n_{2\cdot}}{n_{\cdot 1} - i} \phi_2^i \quad = \quad \alpha/2$$

When the odds ratio equals zero, which occurs when either $n_{11} = 0$ or $n_{22} = 0$, PROC FREQ sets the lower exact confidence limit to zero and determines the upper limit with level α. Similarly, when the odds ratio equals infinity, which occurs when either $n_{12} = 0$ or $n_{21} = 0$, PROC FREQ sets the upper exact confidence limit to infinity and determines the lower limit with level α.

Relative Risks (Cohort Studies)

These measures of relative risk are useful in *cohort* (prospective) study designs, where two samples are identified based on the presence or absence of an explanatory factor. The two samples are observed in future time for the binary (yes-no) response variable under study. Relative risk measures are also useful in cross-sectional studies, where two variables are observed simultaneously. See Stokes, Davis, and Koch (2000) and Agresti (2007) for more information.

The column 1 relative risk is the ratio of the column 1 risk for row 1 to row 2. The column 1 risk for row 1 is the proportion of the row 1 observations classified in column 1,

$$p_1 = n_{11} \;/\; n_{1\cdot}$$

Similarly, the column 1 risk for row 2 is

$$p_2 = n_{21} \;/\; n_{2\cdot}$$

The column 1 relative risk is then computed as

$$RR_1 = p_1 \;/\; p_2$$

A relative risk greater than 1 indicates that the probability of positive response is greater in row 1 than in row 2. Similarly, a relative risk less than 1 indicates that the probability of positive response is less in row 1 than in row 2. The strength of association increases with the deviation from 1.

Asymptotic $100(1 - \alpha)\%$ confidence limits for the column 1 relative risk are computed as

$$\left(RR_1 \times \exp(-z\sqrt{v}),\;\; RR_1 \times \exp(z\sqrt{v}) \right)$$

where

$$v = \text{var}(\ln RR_1) = \big((1-p_1)/n_{11}\big) \; + \; \big((1-p_2)/n_{21}\big)$$

and z is the $100(1-\alpha/2)$th percentile of the standard normal distribution. If either n_{11} or n_{21} is zero, the estimates are not computed.

PROC FREQ computes the column 2 relative risks in the same way.

Exact Unconditional Confidence Limits for the Relative Risk If you specify the RELRISK option in the EXACT statement, PROC FREQ provides exact unconditional confidence limits for the relative risk. PROC FREQ computes the confidence limits by inverting two separate one-sided tests (tail method), where the size of each test is at most $\alpha/2$ and the confidence coefficient is at least $(1-\alpha)$. Exact conditional methods, described in the section "Exact Statistics" on page 2366, do not apply to the relative risk due to the presence of a nuisance parameter (Agresti 1992). The unconditional approach eliminates the nuisance parameter by maximizing the p-value over all possible values of the parameter (Santner and Snell 1980).

By default, PROC FREQ uses the unstandardized relative risk as the test statistic in the confidence limit computations. If you specify the RELRISK(METHOD=FMSCORE) option, the procedure uses the Farrington-Manning relative risk score statistic (Chan and Zhang 1999). The score statistic is a less discrete statistic than the raw relative risk and produces less conservative confidence limits (Agresti and Min 2001). See also Santner et al. (2007) for comparisons of methods for computing exact confidence limits.

See the section "Exact Unconditional Confidence Limits for the Risk Difference" on page 2345 for a description of the method that PROC FREQ uses to compute confidence limits for the relative risk. The test statistic for the relative risk computation is either the unstandardized relative risk (by default) or the relative risk score statistic (if you specify the RELRISK(METHOD=FMSCORE) option). PROC FREQ uses the following form of the unstandardized relative risk, which adds 0.05 to each frequency, to ensure that the statistic is defined when there are zero table cells (Gart and Nam 1988):

$$\hat{rr} = \frac{(n_{11}+0.5) \; / \; (n_{1\cdot}+0.5)}{(n_{21}+0.5) \; / \; (n_{2\cdot}+0.5)}$$

If you specify the RELRISK(METHOD=FMSCORE) option, PROC FREQ uses the relative risk score statistic (Farrington and Manning 1990; Miettinen and Nurminen 1985). This test statistic is computed as

$$z = (\hat{p}_1 - R_0\hat{p}_2) \; / \; \text{se}(\hat{rr})$$

where

$$\text{se}(\hat{rr}) = \sqrt{\tilde{p}_1(1-\tilde{p}_1)/n_{1\cdot} \; + \; R_0^2 \; \tilde{p}_2(1-\tilde{p}_2)/n_{2\cdot}}$$

where $\tilde{p}_1$ and $\tilde{p}_2$ are the maximum likelihood estimators of p_1 and p_2 under the null hypothesis that the relative risk equals R_0. From Farrington and Manning (1990), the maximum likelihood solution is

$$\tilde{p}_1 = (-b - \sqrt{b^2 - 4ac})/2a \quad \text{and} \quad \tilde{p}_2 = \tilde{p}_1/R_0$$

where

$$\begin{aligned} a &= 1+\theta \\ b &= -(R_0(1+\theta\hat{p}_2)+\theta+\hat{p}_1) \\ c &= R_0(\hat{p}_1+\theta\hat{p}_2) \\ \theta &= n_{2\cdot}/n_{1\cdot} \end{aligned}$$

Cochran-Armitage Test for Trend

The TREND option in the TABLES statement provides the Cochran-Armitage test for trend, which tests for trend in binomial proportions across levels of a single factor or covariate. This test is appropriate for a two-way table where one variable has two levels and the other variable is ordinal. The two-level variable represents the response, and the other variable represents an explanatory variable with ordered levels. When the two-way has two columns and R rows, PROC FREQ tests for trend across the R levels of the row variable, and the binomial proportion is computed as the proportion of observations in the first column. When the table has two rows and C columns, PROC FREQ tests for trend across the C levels of the column variable, and the binomial proportion is computed as the proportion of observations in the first row.

The trend test is based on the regression coefficient for the weighted linear regression of the binomial proportions on the scores of the explanatory variable levels. See Margolin (1988) and Agresti (2002) for details. If the table has two columns and R rows, the trend test statistic is computed as

$$T = \sum_{i=1}^{R} n_{i1}(R_i - \bar{R}) \;/\; \sqrt{p_{\cdot 1}\,(1-p_{\cdot 1})\,s^2}$$

where R_i is the score of row i, $\bar{R}$ is the average row score, and

$$s^2 = \sum_{i=1}^{R} n_{i\cdot}(R_i - \bar{R})^2$$

The SCORES= option in the TABLES statement determines the type of row scores used in computing the trend test (and other score-based statistics). The default is SCORES=TABLE. See the section "Scores" on page 2314 for details. For character variables, the table scores for the row variable are the row numbers (for example, 1 for the first row, 2 for the second row, and so on). For numeric variables, the table score for each row is the numeric value of the row level. When you perform the trend test, the explanatory variable might be numeric (for example, dose of a test substance), and the variable values might be appropriate scores. If the explanatory variable has ordinal levels that are not numeric, you can assign meaningful scores to the variable levels. Sometimes equidistant scores, such as the table scores for a character variable, might be appropriate. For more information on choosing scores for the trend test, see Margolin (1988).

The null hypothesis for the Cochran-Armitage test is no trend, which means that the binomial proportion $p_{i1} = n_{i1}/n_{i\cdot}$ is the same for all levels of the explanatory variable. Under the null hypothesis, the trend statistic has an asymptotic standard normal distribution.

PROC FREQ computes one-sided and two-sided p-values for the trend test. When the test statistic is greater than its null hypothesis expected value of zero, PROC FREQ displays the right-sided p-value, which is the

probability of a larger value of the statistic occurring under the null hypothesis. A small right-sided p-value supports the alternative hypothesis of increasing trend in proportions from row 1 to row R. When the test statistic is less than or equal to zero, PROC FREQ displays the left-sided p-value. A small left-sided p-value supports the alternative of decreasing trend.

The one-sided p-value for the trend test is computed as

$$P_1 = \begin{cases} \text{Prob}(Z > T) & \text{if } T > 0 \\ \text{Prob}(Z < T) & \text{if } T \leq 0 \end{cases}$$

where Z has a standard normal distribution. The two-sided p-value is computed as

$$P_2 = \text{Prob}(|Z| > |T|)$$

PROC FREQ also provides exact p-values for the Cochran-Armitage trend test. You can request the exact test by specifying the TREND option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Jonckheere-Terpstra Test

The JT option in the TABLES statement provides the Jonckheere-Terpstra test, which is a nonparametric test for ordered differences among classes. It tests the null hypothesis that the distribution of the response variable does not differ among classes. It is designed to detect alternatives of ordered class differences, which can be expressed as $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_R$ (or $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_R$), with at least one of the inequalities being strict, where τ_i denotes the effect of class i. For such ordered alternatives, the Jonckheere-Terpstra test can be preferable to tests of more general class difference alternatives, such as the Kruskal–Wallis test (produced by the WILCOXON option in the NPAR1WAY procedure). See Pirie (1983) and Hollander and Wolfe (1999) for more information about the Jonckheere-Terpstra test.

The Jonckheere-Terpstra test is appropriate for a two-way table in which an ordinal column variable represents the response. The row variable, which can be nominal or ordinal, represents the classification variable. The levels of the row variable should be ordered according to the ordering you want the test to detect. The order of variable levels is determined by the ORDER= option in the PROC FREQ statement. The default is ORDER=INTERNAL, which orders by unformatted values. If you specify ORDER=DATA, PROC FREQ orders values according to their order in the input data set. For more information about how to order variable levels, see the ORDER= option.

The Jonckheere-Terpstra test statistic is computed by first forming $R(R-1)/2$ Mann-Whitney counts $M_{i,i'}$, where $i < i'$, for pairs of rows in the contingency table,

$$\begin{aligned} M_{i,i'} &= \{ \text{ number of times } X_{i,j} < X_{i',j'}, \quad j = 1,\ldots,n_{i.}; \quad j' = 1,\ldots,n_{i'.} \} \\ &\quad + \tfrac{1}{2} \{ \text{ number of times } X_{i,j} = X_{i',j'}, \quad j = 1,\ldots,n_{i.}; \quad j' = 1,\ldots,n_{i'.} \} \end{aligned}$$

where $X_{i,j}$ is response j in row i. The Jonckheere-Terpstra test statistic is computed as

$$J = \sum_{1 \leq i < i' \leq R} \sum M_{i,i'}$$

This test rejects the null hypothesis of no difference among classes for large values of J. Asymptotic p-values for the Jonckheere-Terpstra test are obtained by using the normal approximation for the distribution of the standardized test statistic. The standardized test statistic is computed as

$$J^* = (J - \mathrm{E}_0(J)) \ / \ \sqrt{\mathrm{var}_0(J)}$$

where $\mathrm{E}_0(J)$ and $\mathrm{var}_0(J)$ are the expected value and variance of the test statistic under the null hypothesis,

$$\mathrm{E}_0(J) = \left(n^2 - \sum_i n_{i\cdot}^2 \right) / 4$$

$$\mathrm{var}_0(J) = A/72 + B/\left(36n(n-1)(n-2)\right) + C/\left(8n(n-1)\right)$$

where

$$A = n(n-1)(2n+5) - \sum_i n_{i\cdot}(n_{i\cdot}-1)(2n_{i\cdot}+5) - \sum_j n_{\cdot j}(n_{\cdot j}-1)(2n_{\cdot j}+5)$$

$$B = \left(\sum_i n_{i\cdot}(n_{i\cdot}-1)(n_{i\cdot}-2) \right) \left(\sum_j n_{\cdot j}(n_{\cdot j}-1)(n_{\cdot j}-2) \right)$$

$$C = \left(\sum_i n_{i\cdot}(n_{i\cdot}-1) \right) \left(\sum_j n_{\cdot j}(n_{\cdot j}-1) \right)$$

PROC FREQ computes one-sided and two-sided p-values for the Jonckheere-Terpstra test. When the standardized test statistic is greater than its null hypothesis expected value of zero, PROC FREQ displays the right-sided p-value, which is the probability of a larger value of the statistic occurring under the null hypothesis. A small right-sided p-value supports the alternative hypothesis of increasing order from row 1 to row R. When the standardized test statistic is less than or equal to zero, PROC FREQ displays the left-sided p-value. A small left-sided p-value supports the alternative of decreasing order from row 1 to row R.

The one-sided p-value for the Jonckheere-Terpstra test, P_1, is computed as

$$P_1 = \begin{cases} \mathrm{Prob}(Z > J^*) & \text{if } J^* > 0 \\ \mathrm{Prob}(Z < J^*) & \text{if } J^* \leq 0 \end{cases}$$

where Z has a standard normal distribution. The two-sided p-value, P_2, is computed as

$$P_2 = \mathrm{Prob}(|Z| > |J^*|)$$

PROC FREQ also provides exact p-values for the Jonckheere-Terpstra test. You can request the exact test by specifying the JT option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Tests and Measures of Agreement

When you specify the AGREE option in the TABLES statement, PROC FREQ computes tests and measures of agreement for square tables (that is, for tables where the number of rows equals the number of columns). For two-way tables, these tests and measures include McNemar's test for 2×2 tables, Bowker's test of symmetry, the simple kappa coefficient, and the weighted kappa coefficient. For multiple strata (n-way tables, where $n > 2$), PROC FREQ also computes the overall simple kappa coefficient and the overall weighted kappa coefficient, as well as tests for equal kappas (simple and weighted) among strata. Cochran's Q is computed for multiway tables when each variable has two levels, that is, for $h \times 2 \times 2$ tables.

PROC FREQ computes the kappa coefficients (simple and weighted), their asymptotic standard errors, and their confidence limits when you specify the AGREE option in the TABLES statement. If you also specify the KAPPA option in the TEST statement, then PROC FREQ computes the asymptotic test of the hypothesis that simple kappa equals zero. Similarly, if you specify the WTKAP option in the TEST statement, PROC FREQ computes the asymptotic test for weighted kappa.

In addition to the asymptotic tests described in this section, PROC FREQ provides exact p-values for McNemar's test, the simple kappa coefficient test, and the weighted kappa coefficient test. You can request these exact tests by specifying the corresponding options in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

The following sections provide the formulas that PROC FREQ uses to compute the AGREE statistics. For information about the use and interpretation of these statistics, see Agresti (2002), Agresti (2007), Fleiss, Levin, and Paik (2003), and the other references cited for each statistic.

McNemar's Test

PROC FREQ computes McNemar's test for 2×2 tables when you specify the AGREE option. McNemar's test is appropriate when you are analyzing data from matched pairs of subjects with a dichotomous (yes-no) response. It tests the null hypothesis of marginal homogeneity, or $p_{1\cdot} = p_{\cdot 1}$. McNemar's test is computed as

$$Q_M = (n_{12} - n_{21})^2 \; / \; (n_{12} + n_{21})$$

Under the null hypothesis, Q_M has an asymptotic chi-square distribution with one degree of freedom. See McNemar (1947), as well as the general references cited in the preceding section. In addition to the asymptotic test, PROC FREQ also computes the exact p-value for McNemar's test when you specify the MCNEM option in the EXACT statement.

Bowker's Test of Symmetry

For Bowker's test of symmetry, the null hypothesis is that the cell proportions are symmetric, or that $p_{ij} = p_{ji}$ for all pairs of table cells. For 2×2 tables, Bowker's test is identical to McNemar's test, and so PROC FREQ provides Bowker's test for square tables larger than 2×2.

Bowker's test of symmetry is computed as

$$Q_B = \sum_{i<j}\sum (n_{ij} - n_{ji})^2 \;/\; (n_{ij} + n_{ji})$$

For large samples, Q_B has an asymptotic chi-square distribution with $R(R-1)/2$ degrees of freedom under the null hypothesis of symmetry. See Bowker (1948) for details.

Simple Kappa Coefficient

The simple kappa coefficient, introduced by Cohen (1960), is a measure of interrater agreement. PROC FREQ computes the simple kappa coefficient as

$$\hat{\kappa} = (P_o - P_e) \;/\; (1 - P_e)$$

where $P_o = \sum_i p_{ii}$ and $P_e = \sum_i p_{i\cdot}p_{\cdot i}$. If the two response variables are viewed as two independent ratings of the n subjects, the kappa coefficient equals +1 when there is complete agreement of the raters. When the observed agreement exceeds chance agreement, kappa is positive, with its magnitude reflecting the strength of agreement. Although this is unusual in practice, kappa is negative when the observed agreement is less than chance agreement. The minimum value of kappa is between -1 and 0, depending on the marginal proportions.

The asymptotic variance of the simple kappa coefficient is computed as

$$\mathrm{var}(\hat{\kappa}) = (A + B - C) \;/\; (1 - P_e)^2\, n$$

where

$$\begin{aligned} A &= \sum_i p_{ii}\,(1 - (p_{i\cdot} + p_{\cdot i})(1 - \hat{\kappa}))^2 \\ B &= (1-\hat{\kappa})^2 \sum_{i\neq j}\sum p_{ij}(p_{\cdot i} + p_{j\cdot})^2 \\ C &= (\,\hat{\kappa} - P_e(1-\hat{\kappa})\,)^2 \end{aligned}$$

See Fleiss, Cohen, and Everitt (1969) for details.

PROC FREQ computes confidence limits for the simple kappa coefficient as

$$\hat{\kappa} \;\pm\; \left(\, z_{\alpha/2} \times \sqrt{\mathrm{var}(\hat{\kappa})}\,\right)$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution. The value of α is determined by the value of the ALPHA= option, which, by default, equals 0.05 and produces 95% confidence limits.

To compute an asymptotic test for the kappa coefficient, PROC FREQ uses the standardized test statistic $\hat{\kappa}^*$, which has an asymptotic standard normal distribution under the null hypothesis that kappa equals zero. The standardized test statistic is computed as

$$\hat{\kappa}^* = \hat{\kappa} \;/\; \sqrt{\mathrm{var}_0(\hat{\kappa})}$$

where $\text{var}_0(\hat{\kappa})$ is the variance of the kappa coefficient under the null hypothesis,

$$\text{var}_0(\hat{\kappa}) = \left(P_e + P_e^2 - \sum_i p_{i\cdot} p_{\cdot i} (p_{i\cdot} + p_{\cdot i}) \right) \Big/ (1 - P_e)^2 \, n$$

See Fleiss, Levin, and Paik (2003) for details.

PROC FREQ also provides an exact test for the simple kappa coefficient. You can request the exact test by specifying the KAPPA or AGREE option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Weighted Kappa Coefficient

The weighted kappa coefficient is a generalization of the simple kappa coefficient that uses weights to quantify the relative difference between categories. For 2×2 tables, the weighted kappa coefficient equals the simple kappa coefficient. PROC FREQ displays the weighted kappa coefficient only for tables larger than 2×2. PROC FREQ computes the kappa weights from the column scores, by using either Cicchetti-Allison weights or Fleiss-Cohen weights, both of which are described in the following section. The weights w_{ij} are constructed so that $0 \le w_{ij} < 1$ for all $i \ne j$, $w_{ii} = 1$ for all i, and $w_{ij} = w_{ji}$. The weighted kappa coefficient is computed as

$$\hat{\kappa}_w = \left(P_{o(w)} - P_{e(w)} \right) \Big/ \left(1 - P_{e(w)} \right)$$

where

$$P_{o(w)} = \sum_i \sum_j w_{ij} \, p_{ij}$$

$$P_{e(w)} = \sum_i \sum_j w_{ij} \, p_{i\cdot} \, p_{\cdot j}$$

The asymptotic variance of the weighted kappa coefficient is

$$\text{var}(\hat{\kappa}_w) = \left(\sum_i \sum_j p_{ij} \left(w_{ij} - (\overline{w}_{i\cdot} + \overline{w}_{\cdot j})(1 - \hat{\kappa}_w) \right)^2 - \left(\hat{\kappa}_w - P_{e(w)}(1 - \hat{\kappa}_w) \right)^2 \right) \Big/ (1 - P_{e(w)})^2 \, n$$

where

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

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

See Fleiss, Cohen, and Everitt (1969) for details.

PROC FREQ computes confidence limits for the weighted kappa coefficient as

$$\hat{\kappa}_w \pm \left(z_{\alpha/2} \times \sqrt{\text{var}(\hat{\kappa}_w)} \right)$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution. The value of α is determined by the value of the ALPHA= option, which, by default, equals 0.05 and produces 95% confidence limits.

To compute an asymptotic test for the weighted kappa coefficient, PROC FREQ uses the standardized test statistic $\hat{\kappa}_w^*$, which has an asymptotic standard normal distribution under the null hypothesis that weighted kappa equals zero. The standardized test statistic is computed as

$$\hat{\kappa}_w^* = \hat{\kappa}_w \;/\; \sqrt{\mathrm{var}_0(\hat{\kappa}_w)}$$

where $\mathrm{var}_0(\hat{\kappa}_w)$ is the variance of the weighted kappa coefficient under the null hypothesis,

$$\mathrm{var}_0(\hat{\kappa}_w) = \left(\sum_i \sum_j p_{i\cdot} p_{\cdot j} \left(w_{ij} - (\overline{w}_{i\cdot} + \overline{w}_{\cdot j}) \right)^2 - P_{e(w)}^2 \right) \;/\; (1 - P_{e(w)})^2 \, n$$

See Fleiss, Levin, and Paik (2003) for details.

PROC FREQ also provides an exact test for the weighted kappa coefficient. You can request the exact test by specifying the WTKAPPA or AGREE option in the EXACT statement. See the section "Exact Statistics" on page 2366 for more information.

Weights PROC FREQ computes kappa coefficient weights by using the column scores and one of the two available weight types. The column scores are determined by the SCORES= option in the TABLES statement. The two available types of kappa weights are Cicchetti-Allison and Fleiss-Cohen weights. By default, PROC FREQ uses Cicchetti-Allison weights. If you specify (WT=FC) with the AGREE option, then PROC FREQ uses Fleiss-Cohen weights to compute the weighted kappa coefficient.

PROC FREQ computes Cicchetti-Allison kappa coefficient weights as

$$w_{ij} = 1 - \frac{|C_i - C_j|}{C_C - C_1}$$

where C_i is the score for column i and C is the number of categories or columns. See Cicchetti and Allison (1971) for details.

The SCORES= option in the TABLES statement determines the type of column scores used to compute the kappa weights (and other score-based statistics). The default is SCORES=TABLE. See the section "Scores" on page 2314 for details. For numeric variables, table scores are the values of the variable levels. You can assign numeric values to the levels in a way that reflects their level of similarity. For example, suppose you have four levels and order them according to similarity. If you assign them values of 0, 2, 4, and 10, the Cicchetti-Allison kappa weights take the following values: $w_{12} = 0.8$, $w_{13} = 0.6$, $w_{14} = 0$, $w_{23} = 0.8$, $w_{24} = 0.2$, and $w_{34} = 0.4$. Note that when there are only two categories (that is, $C = 2$), the weighted kappa coefficient is identical to the simple kappa coefficient.

If you specify (WT=FC) with the AGREE option in the TABLES statement, PROC FREQ computes Fleiss-Cohen kappa coefficient weights as

$$w_{ij} = 1 - \frac{(C_i - C_j)^2}{(C_C - C_1)^2}$$

See Fleiss and Cohen (1973) for details.

For the preceding example, the Fleiss-Cohen kappa weights are: $w_{12} = 0.96$, $w_{13} = 0.84$, $w_{14} = 0$, $w_{23} = 0.96$, $w_{24} = 0.36$, and $w_{34} = 0.64$.

Overall Kappa Coefficient

When there are multiple strata, PROC FREQ combines the stratum-level estimates of kappa into an overall estimate of the supposed common value of kappa. Assume there are q strata, indexed by $h = 1, 2, \ldots, q$, and let $\mathrm{var}(\hat{\kappa}_h)$ denote the variance of $\hat{\kappa}_h$. The estimate of the overall kappa coefficient is computed as

$$\hat{\kappa}_T = \sum_{h=1}^{q} \frac{\hat{\kappa}_h}{\mathrm{var}(\hat{\kappa}_h)} \Big/ \sum_{h=1}^{q} \frac{1}{\mathrm{var}(\hat{\kappa}_h)}$$

See Fleiss, Levin, and Paik (2003) for details.

PROC FREQ computes an estimate of the overall weighted kappa in the same way.

Tests for Equal Kappa Coefficients

When there are multiple strata, the following chi-square statistic tests whether the stratum-level values of kappa are equal:

$$Q_K = \sum_{h=1}^{q} (\hat{\kappa}_h - \hat{\kappa}_T)^2 \Big/ \mathrm{var}(\hat{\kappa}_h)$$

Under the null hypothesis of equal kappas for the q strata, Q_K has an asymptotic chi-square distribution with $q - 1$ degrees of freedom. See Fleiss, Levin, and Paik (2003) for more information. PROC FREQ computes a test for equal weighted kappa coefficients in the same way.

Cochran's Q Test

Cochran's Q is computed for multiway tables when each variable has two levels, that is, for $2 \times 2 \cdots \times 2$ tables. Cochran's Q statistic is used to test the homogeneity of the one-dimensional margins. Let m denote the number of variables and N denote the total number of subjects. Cochran's Q statistic is computed as

$$Q_C = m(m-1) \left(\sum_{j=1}^{m} T_j^2 - T^2 \right) \Big/ \left(mT - \sum_{k=1}^{N} S_k^2 \right)$$

where T_j is the number of positive responses for variable j, T is the total number of positive responses over all variables, and S_k is the number of positive responses for subject k. Under the null hypothesis, Cochran's Q has an asymptotic chi-square distribution with $m - 1$ degrees of freedom. See Cochran (1950) for details. When there are only two binary response variables ($m = 2$), Cochran's Q simplifies to McNemar's test. When there are more than two response categories, you can test for marginal homogeneity by using the repeated measures capabilities of the CATMOD procedure.

Tables with Zero Rows and Columns

The AGREE statistics are defined only for square tables, where the number of rows equals the number of columns. If the table is not square, PROC FREQ does not compute AGREE statistics. In the kappa statistic framework, where two independent raters assign ratings to each of n subjects, suppose one of the raters does not use all possible r rating levels. If the corresponding table has r rows but only $r - 1$ columns, then

the table is not square and PROC FREQ does not compute AGREE statistics. To create a square table in this situation, use the ZEROS option in the WEIGHT statement, which requests that PROC FREQ include observations with zero weights in the analysis. Include zero-weight observations in the input data set to represent any rating levels that are not used by a rater, so that the input data set has at least one observation for each possible rater and rating combination. The analysis then includes all rating levels, even when all levels are not actually assigned by both raters. The resulting table (of rater 1 by rater 2) is a square table, and AGREE statistics can be computed.

For more information, see the description of the ZEROS option. By default, PROC FREQ does not process observations that have zero weights, because these observations do not contribute to the total frequency count, and because any resulting zero-weight row or column causes many of the tests and measures of association to be undefined. However, kappa statistics are defined for tables with a zero-weight row or column, and the ZEROS option makes it possible to input zero-weight observations and construct the tables needed to compute kappas.

Cochran-Mantel-Haenszel Statistics

The CMH option in the TABLES statement gives a stratified statistical analysis of the relationship between the row and column variables after controlling for the strata variables in a multiway table. For example, for the table request A*B*C*D, the CMH option provides an analysis of the relationship between C and D, after controlling for A and B. The stratified analysis provides a way to adjust for the possible confounding effects of A and B without being forced to estimate parameters for them.

The CMH analysis produces Cochran-Mantel-Haenszel statistics, which include the correlation statistic, the ANOVA (row mean scores) statistic, and the general association statistic. For 2×2 tables, the CMH option also provides Mantel-Haenszel and logit estimates of the common odds ratio and the common relative risks, as well as the Breslow-Day test for homogeneity of the odds ratios.

Exact statistics are also available for stratified 2×2 tables. If you specify the EQOR option in the EXACT statement, PROC FREQ provides Zelen's exact test for equal odds ratios. If you specify the COMOR option in the EXACT statement, PROC FREQ provides exact confidence limits for the common odds ratio and an exact test that the common odds ratio equals one.

Let the number of strata be denoted by q, indexing the strata by $h = 1, 2, \ldots, q$. Each stratum contains a contingency table with X representing the row variable and Y representing the column variable. For table h, denote the cell frequency in row i and column j by n_{hij}, with corresponding row and column marginal totals denoted by $n_{hi\cdot}$ and $n_{h\cdot j}$, and the overall stratum total by n_h.

Because the formulas for the Cochran-Mantel-Haenszel statistics are more easily defined in terms of matrices, the following notation is used. Vectors are presumed to be column vectors unless they are transposed ($'$).

$$
\begin{aligned}
\mathbf{n}'_{hi} &= (n_{hi1}, n_{hi2}, \ldots, n_{hiC}) && (1 \times C) \\
\mathbf{n}'_{h} &= (\mathbf{n}'_{h1}, \mathbf{n}'_{h2}, \ldots, \mathbf{n}'_{hR}) && (1 \times RC) \\
p_{hi\cdot} &= n_{hi\cdot} \,/\, n_h && (1 \times 1) \\
p_{h\cdot j} &= n_{h\cdot j} \,/\, n_h && (1 \times 1) \\
\mathbf{P}'_{h*\cdot} &= (p_{h1\cdot}, p_{h2\cdot}, \ldots, p_{hR\cdot}) && (1 \times R) \\
\mathbf{P}'_{h\cdot *} &= (p_{h\cdot 1}, p_{h\cdot 2}, \ldots, p_{h\cdot C}) && (1 \times C)
\end{aligned}
$$

Assume that the strata are independent and that the marginal totals of each stratum are fixed. The null hypothesis, H_0, is that there is no association between X and Y in any of the strata. The corresponding model is the multiple hypergeometric; this implies that, under H_0, the expected value and covariance matrix of the frequencies are, respectively,

$$\mathbf{m}_h = \mathbf{E}[\mathbf{n}_h \mid H_0] = n_h(\mathbf{P}_{h\cdot *} \otimes \mathbf{P}_{h*\cdot})$$

$$\mathbf{var}[\mathbf{n}_h \mid H_0] = c \left((\mathbf{D}_{\mathbf{P}h\cdot *} - \mathbf{P}_{h\cdot *}\mathbf{P}'_{h\cdot *}) \otimes (\mathbf{D}_{\mathbf{P}h*\cdot} - \mathbf{P}_{h*\cdot}\mathbf{P}'_{h*\cdot}) \right)$$

where

$$c = n_h^2 \,/\, (n_h - 1)$$

and where $\otimes$ denotes Kronecker product multiplication and $\mathbf{D}_\mathbf{a}$ is a diagonal matrix with the elements of $\mathbf{a}$ on the main diagonal.

The generalized CMH statistic (Landis, Heyman, and Koch 1978) is defined as

$$Q_{CMH} = \mathbf{G}'{\mathbf{V}_\mathbf{G}}^{-1}\mathbf{G}$$

where

$$\mathbf{G} = \sum_h \mathbf{B}_h(\mathbf{n}_h - \mathbf{m}_h)$$

$$\mathbf{V}_\mathbf{G} = \sum_h \mathbf{B}_h \left(\mathit{Var}(\mathbf{n}_h \mid H_0)\right) \mathbf{B}'_h$$

and where

$$\mathbf{B}_h = \mathbf{C}_h \otimes \mathbf{R}_h$$

is a matrix of fixed constants based on column scores $\mathbf{C}_h$ and row scores $\mathbf{R}_h$. When the null hypothesis is true, the CMH statistic has an asymptotic chi-square distribution with degrees of freedom equal to the rank of $\mathbf{B}_h$. If $\mathbf{V}_\mathbf{G}$ is found to be singular, PROC FREQ prints a message and sets the value of the CMH statistic to missing.

PROC FREQ computes three CMH statistics by using this formula for the generalized CMH statistic, with different row and column score definitions for each statistic. The CMH statistics that PROC FREQ computes are the correlation statistic, the ANOVA (row mean scores) statistic, and the general association statistic. These statistics test the null hypothesis of no association against different alternative hypotheses. The following sections describe the computation of these CMH statistics.

CAUTION: The CMH statistics have low power for detecting an association in which the patterns of association for some of the strata are in the opposite direction of the patterns displayed by other strata. Thus, a nonsignificant CMH statistic suggests either that there is no association or that no pattern of association has enough strength or consistency to dominate any other pattern.

Correlation Statistic

The correlation statistic, popularized by Mantel and Haenszel (1959) and Mantel (1963), has one degree of freedom and is known as the Mantel-Haenszel statistic.

The alternative hypothesis for the correlation statistic is that there is a linear association between X and Y in at least one stratum. If either X or Y does not lie on an ordinal (or interval) scale, then this statistic is not meaningful.

To compute the correlation statistic, PROC FREQ uses the formula for the generalized CMH statistic with the row and column scores determined by the SCORES= option in the TABLES statement. See the section "Scores" on page 2314 for more information about the available score types. The matrix of row scores $\mathbf{R}_h$ has dimension $1 \times R$, and the matrix of column scores $\mathbf{C}_h$ has dimension $1 \times C$.

When there is only one stratum, this CMH statistic reduces to $(n-1)r^2$, where r is the Pearson correlation coefficient between X and Y. When nonparametric (RANK or RIDIT) scores are specified, the statistic reduces to $(n-1)r_s^2$, where r_s is the Spearman rank correlation coefficient between X and Y. When there is more than one stratum, this CMH statistic becomes a stratum-adjusted correlation statistic.

ANOVA (Row Mean Scores) Statistic

The ANOVA statistic can be used only when the column variable Y lies on an ordinal (or interval) scale so that the mean score of Y is meaningful. For the ANOVA statistic, the mean score is computed for each row of the table, and the alternative hypothesis is that, for at least one stratum, the mean scores of the R rows are unequal. In other words, the statistic is sensitive to location differences among the R distributions of Y.

The matrix of column scores $\mathbf{C}_h$ has dimension $1 \times C$, and the column scores are determined by the SCORES= option.

The matrix of row scores $\mathbf{R}_h$ has dimension $(R-1) \times R$ and is created internally by PROC FREQ as

$$\mathbf{R}_h = [\mathbf{I}_{R-1}, -\mathbf{J}_{R-1}]$$

where $\mathbf{I}_{R-1}$ is an identity matrix of rank $R-1$ and $\mathbf{J}_{R-1}$ is an $(R-1) \times 1$ vector of ones. This matrix has the effect of forming $R-1$ independent contrasts of the R mean scores.

When there is only one stratum, this CMH statistic is essentially an analysis of variance (ANOVA) statistic in the sense that it is a function of the variance ratio F statistic that would be obtained from a one-way ANOVA on the dependent variable Y. If nonparametric scores are specified in this case, then the ANOVA statistic is a Kruskal-Wallis test.

If there is more than one stratum, then this CMH statistic corresponds to a stratum-adjusted ANOVA or Kruskal-Wallis test. In the special case where there is one subject per row and one subject per column in the contingency table of each stratum, this CMH statistic is identical to Friedman's chi-square. See Example 36.9 for an illustration.

General Association Statistic

The alternative hypothesis for the general association statistic is that, for at least one stratum, there is some kind of association between X and Y. This statistic is always interpretable because it does not require an ordinal scale for either X or Y.

For the general association statistic, the matrix $\mathbf{R}_h$ is the same as the one used for the ANOVA statistic. The matrix $\mathbf{C}_h$ is defined similarly as

$$\mathbf{C}_h = [\mathbf{I}_{C-1}, -\mathbf{J}_{C-1}]$$

PROC FREQ generates both score matrices internally. When there is only one stratum, then the general association CMH statistic reduces to $Q_P(n-1)/n$, where Q_P is the Pearson chi-square statistic. When there is more than one stratum, then the CMH statistic becomes a stratum-adjusted Pearson chi-square statistic. Note that a similar adjustment can be made by summing the Pearson chi-squares across the strata. However, the latter statistic requires a large sample size in each stratum to support the resulting chi-square distribution with $q(R-1)(C-1)$ degrees of freedom. The CMH statistic requires only a large overall sample size because it has only $(R-1)(C-1)$ degrees of freedom.

See Cochran (1954); Mantel and Haenszel (1959); Mantel (1963); Birch (1965); and Landis, Heyman, and Koch (1978).

Mantel-Fleiss Criterion

If you specify the CMH(MANTELFLEISS) option in the TABLES statement, PROC FREQ computes the Mantel-Fleiss criterion for stratified 2×2 tables. The Mantel-Fleiss criterion can be used to assess the validity of the chi-square approximation for the distribution of the Mantel-Haenszel statistic for 2×2 tables. See Mantel and Fleiss (1980), Mantel and Haenszel (1959), Stokes, Davis, and Koch (2000), and Dimitrienko et al. (2005) for details.

The Mantel-Fleiss criterion is computed as

$$MF = \min\left(\left[\sum_h m_{h11} - \sum_h (n_{h11})_L\right], \left[\sum_h (n_{h11})_U - \sum_h m_{h11}\right]\right)$$

where m_{h11} is the expected value of n_{h11} under the hypothesis of no association between the row and column variables in table h, $(n_{h11})_L$ is the minimum possible value of the table cell frequency, and $(n_{h11})_U$ is the maximum possible value,

$$\begin{aligned} m_{h11} &= n_{h1\cdot}\, n_{h\cdot 1} \,/\, n_h \\ (n_{h11})_L &= \max(\,0,\ \ n_{h1\cdot} - n_{h\cdot 2}\,) \\ (n_{h11})_U &= \min(\,n_{h\cdot 1},\ \ n_{h1\cdot}\,) \end{aligned}$$

The Mantel-Fleiss guideline accepts the validity of the Mantel-Haenszel approximation when the value of the criterion is at least 5. When the criterion is less than 5, PROC FREQ displays a warning.

Adjusted Odds Ratio and Relative Risk Estimates

The CMH option provides adjusted odds ratio and relative risk estimates for stratified 2×2 tables. For each of these measures, PROC FREQ computes a Mantel-Haenszel estimate and a logit estimate. These estimates apply to n-way table requests in the TABLES statement, when the row and column variables both have two levels.

For example, for the table request A*B*C*D, if the row and column variables C and D both have two levels, PROC FREQ provides odds ratio and relative risk estimates, adjusting for the confounding variables A and B.

The choice of an appropriate measure depends on the study design. For case-control (retrospective) studies, the odds ratio is appropriate. For cohort (prospective) or cross-sectional studies, the relative risk is appropriate. See the section "Odds Ratio and Relative Risks for 2 x 2 Tables" on page 2346 for more information on these measures.

Throughout this section, z denotes the $100(1-\alpha/2)$th percentile of the standard normal distribution.

Odds Ratio, Case-Control Studies PROC FREQ provides Mantel-Haenszel and logit estimates for the common odds ratio for stratified 2×2 tables.

Mantel-Haenszel Estimator The Mantel-Haenszel estimate of the common odds ratio is computed as

$$OR_{MH} = \left(\sum_h n_{h11}\, n_{h22}/n_h \right) \;/\; \left(\sum_h n_{h12}\, n_{h21}/n_h \right)$$

It is always computed unless the denominator is zero. See Mantel and Haenszel (1959) and Agresti (2002) for details.

To compute confidence limits for the common odds ratio, PROC FREQ uses the Greenland and Robins (1985) variance estimate for $\ln(OR_{MH})$. The $100(1-\alpha/2)$ confidence limits for the common odds ratio are

$$\left(\; OR_{MH} \times \exp(-z\hat{\sigma}), \quad OR_{MH} \times \exp(z\hat{\sigma}) \; \right)$$

where

$$\begin{aligned}
\hat{\sigma}^2 &= \widehat{\mathrm{var}}(\, \ln(OR_{MH}) \,) \\
&= \frac{\sum_h (n_{h11} + n_{h22})(n_{h11}\, n_{h22})/n_h^2}{2\left(\sum_h n_{h11}\, n_{h22}/n_h\right)^2} \\
&\quad + \frac{\sum_h [(n_{h11} + n_{h22})(n_{h12}\, n_{h21}) + (n_{h12} + n_{h21})(n_{h11}\, n_{h22})]/n_h^2}{2\left(\sum_h n_{h11}\, n_{h22}/n_h\right)\left(\sum_h n_{h12}\, n_{h21}/n_h\right)} \\
&\quad + \frac{\sum_h (n_{h12} + n_{h21})(n_{h12}\, n_{h21})/n_h^2}{2\left(\sum_h n_{h12}\, n_{h21}/n_h\right)^2}
\end{aligned}$$

Note that the Mantel-Haenszel odds ratio estimator is less sensitive to small n_h than the logit estimator.

Logit Estimator The adjusted logit estimate of the common odds ratio (Woolf 1955) is computed as

$$OR_L = \exp\left(\sum_h w_h \ln(OR_h) \;/\; \sum_h w_h \right)$$

and the corresponding $100(1-\alpha)\%$ confidence limits are

$$\left(\; OR_L \times \exp\left(-z \Big/ \sqrt{\sum_h w_h} \right), \quad OR_L \times \exp\left(z \Big/ \sqrt{\sum_h w_h} \right) \right)$$

where OR_h is the odds ratio for stratum h, and

$$w_h = 1/\text{var}(\ln(OR_h))$$

If any table cell frequency in a stratum h is zero, PROC FREQ adds 0.5 to each cell of the stratum before computing OR_h and w_h (Haldane 1955) for the logit estimate. The procedure prints a warning when this occurs.

Relative Risks, Cohort Studies PROC FREQ provides Mantel-Haenszel and logit estimates of the common relative risks for stratified 2×2 tables.

Mantel-Haenszel Estimator The Mantel-Haenszel estimate of the common relative risk for column 1 is computed as

$$RR_{MH} = \left(\sum_h n_{h11}\, n_{h2\cdot} \,/\, n_h \right) \,/\, \left(\sum_h n_{h21}\, n_{h1\cdot} \,/\, n_h \right)$$

It is always computed unless the denominator is zero. See Mantel and Haenszel (1959) and Agresti (2002) for more information.

To compute confidence limits for the common relative risk, PROC FREQ uses the Greenland and Robins (1985) variance estimate for $\log(RR_{MH})$. The $100(1-\alpha/2)$ confidence limits for the common relative risk are

$$\left(RR_{MH} \times \exp(-z\hat{\sigma}),\quad RR_{MH} \times \exp(z\hat{\sigma}) \right)$$

where

$$\hat{\sigma}^2 = \widehat{\text{var}}(\ln(RR_{MH})) = \frac{\sum_h (n_{h1\cdot}\, n_{h2\cdot}\, n_{h\cdot 1} - n_{h11}\, n_{h21}\, n_h)/n_h^2}{\left(\sum_h n_{h11}\, n_{h2\cdot}/n_h\right)\left(\sum_h n_{h21}\, n_{h1\cdot}/n_h\right)}$$

Logit Estimator The adjusted logit estimate of the common relative risk for column 1 is computed as

$$RR_L = \exp\left(\sum_h w_h \ln(RR_h) \,/\, \sum w_h \right)$$

and the corresponding $100(1-\alpha)\%$ confidence limits are

$$\left(RR_L \times \exp\left(-z \,/\, \sqrt{\sum_h w_h} \right),\quad RR_L \times \exp\left(z \,/\, \sqrt{\sum_h w_h} \right) \right)$$

where RR_h is the column 1 relative risk estimate for stratum h and

$$w_h = 1 \,/\, \text{var}(\ln(RR_h))$$

If n_{h11} or n_{h21} is zero, then PROC FREQ adds 0.5 to each cell of the stratum before computing RR_h and w_h for the logit estimate. The procedure prints a warning when this occurs. See Kleinbaum, Kupper, and Morgenstern (1982, Sections 17.4 and 17.5) for details.

Breslow-Day Test for Homogeneity of the Odds Ratios

When you specify the CMH option, PROC FREQ computes the Breslow-Day test for stratified 2×2 tables. It tests the null hypothesis that the odds ratios for the q strata are equal. When the null hypothesis is true, the statistic has approximately a chi-square distribution with $q - 1$ degrees of freedom. See Breslow and Day (1980) and Agresti (2007) for more information.

The Breslow-Day statistic is computed as

$$Q_{BD} = \sum_h (n_{h11} - \mathrm{E}(n_{h11} \mid OR_{MH}))^2 \ / \ \mathrm{var}(n_{h11} \mid OR_{MH})$$

where E and var denote expected value and variance, respectively. The summation does not include any table with a zero row or column. If OR_{MH} equals zero or if it is undefined, then PROC FREQ does not compute the statistic and prints a warning message.

For the Breslow-Day test to be valid, the sample size should be relatively large in each stratum, and at least 80% of the expected cell counts should be greater than 5. Note that this is a stricter sample size requirement than the requirement for the Cochran-Mantel-Haenszel test for $q \times 2 \times 2$ tables, in that each stratum sample size (not just the overall sample size) must be relatively large. Even when the Breslow-Day test is valid, it might not be very powerful against certain alternatives, as discussed in Breslow and Day (1980).

If you specify the BDT option, PROC FREQ computes the Breslow-Day test with Tarone's adjustment, which subtracts an adjustment factor from Q_{BD} to make the resulting statistic asymptotically chi-square. The Breslow-Day-Tarone statistic is computed as

$$Q_{BDT} = Q_{BD} - \left(\sum_h (n_{h11} - \mathrm{E}(n_{h11} \mid OR_{MH})) \right)^2 \ / \ \sum_h \mathrm{var}(n_{h11} \mid OR_{MH})$$

See Tarone (1985), Jones et al. (1989), and Breslow (1996) for more information.

Zelen's Exact Test for Equal Odds Ratios

If you specify the EQOR option in the EXACT statement, PROC FREQ computes Zelen's exact test for equal odds ratios for stratified 2×2 tables. Zelen's test is an exact counterpart to the Breslow-Day asymptotic test for equal odds ratios. The reference set for Zelen's test includes all possible $q \times 2 \times 2$ tables with the same row, column, and stratum totals as the observed multiway table and with the same sum of cell $(1, 1)$ frequencies as the observed table. The test statistic is the probability of the observed $q \times 2 \times 2$ table conditional on the fixed margins, which is a product of hypergeometric probabilities.

The p-value for Zelen's test is the sum of all table probabilities that are less than or equal to the observed table probability, where the sum is computed over all tables in the reference set determined by the fixed margins and the observed sum of cell $(1, 1)$ frequencies. This test is similar to Fisher's exact test for two-way tables. See Zelen (1971), Hirji (2006), and Agresti (1992) for more information. PROC FREQ computes Zelen's exact test by using the polynomial multiplication algorithm of Hirji et al. (1996).

Exact Confidence Limits for the Common Odds Ratio

If you specify the COMOR option in the EXACT statement, PROC FREQ computes exact confidence limits for the common odds ratio for stratified 2×2 tables. This computation assumes that the odds ratio is constant

over all the 2×2 tables. Exact confidence limits are constructed from the distribution of $S = \sum_h n_{h11}$, conditional on the marginal totals of the 2×2 tables.

Because this is a discrete problem, the confidence coefficient for these exact confidence limits is not exactly $(1 - \alpha)$ but is at least $(1 - \alpha)$. Thus, these confidence limits are conservative. See Agresti (1992) for more information.

PROC FREQ computes exact confidence limits for the common odds ratio by using an algorithm based on Vollset, Hirji, and Elashoff (1991). See also Mehta, Patel, and Gray (1985).

Conditional on the marginal totals of 2×2 table h, let the random variable S_h denote the frequency of table cell $(1, 1)$. Given the row totals $n_{h1\cdot}$ and $n_{h2\cdot}$ and column totals $n_{h\cdot 1}$ and $n_{h\cdot 2}$, the lower and upper bounds for S_h are l_h and u_h,

$$\begin{aligned} l_h &= \max\,(\,0,\;\; n_{h1\cdot} - n_{h\cdot 2}\,) \\ u_h &= \min\,(\,n_{h1\cdot},\;\; n_{h\cdot 1}\,) \end{aligned}$$

Let C_{s_h} denote the hypergeometric coefficient,

$$C_{s_h} = \binom{n_{h\cdot 1}}{s_h} \binom{n_{h\cdot 2}}{n_{h1\cdot} - s_h}$$

and let ϕ denote the common odds ratio. Then the conditional distribution of S_h is

$$P(\, S_h = s_h \mid n_{1\cdot},\; n_{\cdot 1},\; n_{\cdot 2}\,) = C_{s_h}\, \phi^{\,s_h} \,/ \sum_{x\,=\,l_h}^{x\,=\,u_h} C_x\, \phi^{\,x}$$

Summing over all the 2×2 tables, $S = \sum_h S_h$, and the lower and upper bounds of S are l and u,

$$l = \sum_h l_h \quad \text{and} \quad u = \sum_h u_h$$

The conditional distribution of the sum S is

$$P(\, S = s \mid n_{h1\cdot},\; n_{h\cdot 1},\; n_{h\cdot 2};\;\; h = 1, \ldots, q\,) = C_s\, \phi^{\,s} \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x\, \phi^{\,x}$$

where

$$C_s = \sum_{s_1 + \ldots + s_q \,=\, s} \left(\prod_h C_{s_h} \right)$$

Let s_0 denote the observed sum of cell (1,1) frequencies over the q tables. The following two equations are solved iteratively for lower and upper confidence limits for the common odds ratio, ϕ_1 and ϕ_2:

$$\sum_{x\,=\,s_0}^{x\,=\,u} C_x\, {\phi_1}^x \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x\, {\phi_1}^x \;\; = \;\; \alpha/2$$

$$\sum_{x\,=\,l}^{x\,=\,s_0} C_x\, {\phi_2}^x \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x\, {\phi_2}^x \;\; = \;\; \alpha/2$$

When the observed sum s_0 equals the lower bound l, PROC FREQ sets the lower confidence limit to zero and determines the upper limit with level α. Similarly, when the observed sum s_0 equals the upper bound u, PROC FREQ sets the upper confidence limit to infinity and determines the lower limit with level α.

When you specify the COMOR option in the EXACT statement, PROC FREQ also computes the exact test that the common odds ratio equals one. Setting $\phi = 1$, the conditional distribution of the sum S under the null hypothesis becomes

$$P_0(\, S = s \mid n_{h1\cdot},\, n_{h\cdot 1},\, n_{h\cdot 2};\;\; h = 1, \ldots, q\,) = C_s \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x$$

The point probability for this exact test is the probability of the observed sum s_0 under the null hypothesis, conditional on the marginals of the stratified 2×2 tables, and is denoted by $P_0(s_0)$. The expected value of S under the null hypothesis is

$$\mathrm{E}_0(S) = \sum_{x\,=\,l}^{x\,=\,u} x\, C_x \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x$$

The one-sided exact p-value is computed from the conditional distribution as $P_0(S >= s_0)$ or $P_0(S \le s_0)$, depending on whether the observed sum s_0 is greater or less than $\mathrm{E}_0(S)$,

$$P_1 \;=\; P_0(\, S >= s_0\,) = \sum_{x\,=\,s_0}^{x\,=\,u} C_x \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x \quad \text{if} \;\; s_0 > \mathrm{E}_0(S)$$

$$P_1 \;=\; P_0(\, S <= s_0\,) = \sum_{x\,=\,l}^{x\,=\,s_0} C_x \,/ \sum_{x\,=\,l}^{x\,=\,u} C_x \quad \text{if} \;\; s_0 \le \mathrm{E}_0(S)$$

PROC FREQ computes two-sided p-values for this test according to three different definitions. A two-sided p-value is computed as twice the one-sided p-value, setting the result equal to one if it exceeds one,

$$P_2{}^a = 2 \times P_1$$

Additionally, a two-sided p-value is computed as the sum of all probabilities less than or equal to the point probability of the observed sum s_0, summing over all possible values of s, $\;l \le s \le u$,

$$P_2{}^b = \sum_{l \le s \le u : P_0(s) \le P_0(s_0)} P_0(s)$$

Also, a two-sided p-value is computed as the sum of the one-sided p-value and the corresponding area in the opposite tail of the distribution, equidistant from the expected value,

$$P_2{}^c = P_0\,(\,|S - \mathrm{E}_0(S)| \;\ge\; |s_0 - \mathrm{E}_0(S)|\,)$$

Gail-Simon Test for Qualitative Interactions

The GAILSIMON option in the TABLES statement provides the Gail-Simon test for qualitative interaction for stratified 2×2 tables. See Gail and Simon (1985), Silvapulle (2001), and Dimitrienko et al. (2005) for details.

The Gail-Simon test is based on the risk differences in stratified 2×2 tables, where the risk difference is defined as the row 1 risk (proportion in column 1) minus the row 2 risk. See the section "Risks and Risk Differences" on page 2336 for details. By default, the procedure uses column 1 risks to compute the Gail-Simon test. If you specify the GAILSIMON(COLUMN=2) option, the procedure uses column 2 risks.

PROC FREQ computes the Gail-Simon test statistics as described in Gail and Simon (1985),

$$\begin{aligned} Q- &= \sum_h (d_h/s_h)^2 \; I(d_h > 0) \\ Q+ &= \sum_h (d_h/s_h)^2 \; I(d_h < 0) \\ Q &= \min\,(Q-,\; Q+) \end{aligned}$$

where d_h is the risk difference in table h, s_h is the standard error of the risk difference, and $I(d_h > 0)$ equals 1 if $d_h > 0$ and 0 otherwise. Similarly, $I(d_h < 0)$ equals 1 if $d_h < 0$ and 0 otherwise. The q 2×2 tables (strata) are indexed by $h = 1, 2, \ldots, q$.

The p-values for the Gail-Simon statistics are computed as

$$\begin{aligned} p(Q-) &= \sum_h (1 - F_h(Q-))\; B(h; n = q, p = 0.5) \\ p(Q+) &= \sum_h (1 - F_h(Q+))\; B(h; n = q, p = 0.5) \\ p(Q) &= \sum_{h=1}^{q-1} (1 - F_h(Q))\; B(h; n = (q-1), p = 0.5) \end{aligned}$$

where $F_h(\cdot)$ is the cumulative chi-square distribution function with h degrees of freedom and $B(h; n, p)$ is the binomial probability function with parameters n and p. The statistic Q tests the null hypothesis of no qualitative interaction. The statistic $Q-$ tests the null hypothesis of positive risk differences. A small p-value for $Q-$ indicates negative differences; similarly, a small p-value for $Q+$ indicates positive risk differences.

Exact Statistics

Exact statistics can be useful in situations where the asymptotic assumptions are not met, and so the asymptotic p-values are not close approximations for the true p-values. Standard asymptotic methods involve the assumption that the test statistic follows a particular distribution when the sample size is sufficiently large. When the sample size is not large, asymptotic results might not be valid, with the asymptotic p-values differing perhaps substantially from the exact p-values. Asymptotic results might also be unreliable when the distribution of the data is sparse, skewed, or heavily tied. See Agresti (2007) and Bishop, Fienberg, and Holland (1975) for more information. Exact computations are based on the statistical theory of exact conditional inference for contingency tables, reviewed by Agresti (1992).

In addition to computation of exact p-values, PROC FREQ provides the option of estimating exact p-values by Monte Carlo simulation. This can be useful for problems that are so large that exact computations require a great amount of time and memory, but for which asymptotic approximations might not be sufficient.

Exact statistics are available for many PROC FREQ tests. For one-way tables, PROC FREQ provides exact p-values for the binomial proportion tests and the chi-square goodness-of-fit test. Exact (Clopper-Pearson) confidence limits are available for the binomial proportion. For two-way tables, PROC FREQ provides exact p-values for the following tests: Pearson chi-square test, likelihood-ratio chi-square test, Mantel-Haenszel chi-square test, Fisher's exact test, Jonckheere-Terpstra test, and Cochran-Armitage test for trend. PROC FREQ also computes exact p-values for tests of the following statistics: Kendall's tau-b, Stuart's tau-c, Somers' $D(C|R)$, Somers' $D(R|C)$, Pearson correlation coefficient, Spearman correlation coefficient, simple kappa coefficient, and weighted kappa coefficient. For 2×2 tables, PROC FREQ provides McNemar's exact test and exact confidence limits for the odds ratio. PROC FREQ also provides exact unconditional confidence limits for the proportion (risk) difference and for the relative risk. For stratified 2×2 tables, PROC FREQ provides Zelen's exact test for equal odds ratios, exact confidence limits for the common odds ratio, and an exact test for the common odds ratio.

The following sections summarize the exact computational algorithms, define the exact p-values that PROC FREQ computes, discuss the computational resource requirements, and describe the Monte Carlo estimation option.

Computational Algorithms

PROC FREQ computes exact p-values for general $R \times C$ tables by using the network algorithm developed by Mehta and Patel (1983). This algorithm provides a substantial advantage over direct enumeration, which can be very time-consuming and feasible only for small problems. See Agresti (1992) for a review of algorithms for computation of exact p-values, and see Mehta, Patel, and Tsiatis (1984) and Mehta, Patel, and Senchaudhuri (1991) for information about the performance of the network algorithm.

The reference set for a given contingency table is the set of all contingency tables with the observed marginal row and column sums. Corresponding to this reference set, the network algorithm forms a directed acyclic network consisting of nodes in a number of stages. A path through the network corresponds to a distinct table in the reference set. The distances between nodes are defined so that the total distance of a path through the network is the corresponding value of the test statistic. At each node, the algorithm computes the shortest and longest path distances for all the paths that pass through that node. For statistics that can be expressed as a linear combination of cell frequencies multiplied by increasing row and column scores, PROC FREQ computes shortest and longest path distances by using the algorithm of Agresti, Mehta, and Patel (1990). For statistics of other forms, PROC FREQ computes an upper bound for the longest path and a lower bound for the shortest path by following the approach of Valz and Thompson (1994).

The longest and shortest path distances or bounds for a node are compared to the value of the test statistic to determine whether all paths through the node contribute to the p-value, none of the paths through the node contribute to the p-value, or neither of these situations occurs. If all paths through the node contribute, the p-value is incremented accordingly, and these paths are eliminated from further analysis. If no paths contribute, these paths are eliminated from the analysis. Otherwise, the algorithm continues, still processing this node and the associated paths. The algorithm finishes when all nodes have been accounted for.

In applying the network algorithm, PROC FREQ uses full numerical precision to represent all statistics, row and column scores, and other quantities involved in the computations. Although it is possible to use

rounding to improve the speed and memory requirements of the algorithm, PROC FREQ does not do this because it can result in reduced accuracy of the *p*-values.

For one-way tables, PROC FREQ computes the exact chi-square goodness-of-fit test by the method of Radlow and Alf (1975). PROC FREQ generates all possible one-way tables with the observed total sample size and number of categories. For each possible table, PROC FREQ compares its chi-square value with the value for the observed table. If the table's chi-square value is greater than or equal to the observed chi-square, PROC FREQ increments the exact *p*-value by the probability of that table, which is calculated under the null hypothesis by using the multinomial frequency distribution. By default, the null hypothesis states that all categories have equal proportions. If you specify null hypothesis proportions or frequencies by using the TESTP= or TESTF= option in the TABLES statement, then PROC FREQ calculates the exact chi-square test based on that null hypothesis.

Other exact computations are described in sections about the individual statistics. See the section "Binomial Proportion" on page 2329 for details about how PROC FREQ computes exact confidence limits and tests for the binomial proportion. See the section "Odds Ratio and Relative Risks for 2 x 2 Tables" on page 2346 for information about computation of exact confidence limits for the odds ratio for 2×2 tables. Also, see the sections "Exact Unconditional Confidence Limits for the Risk Difference" on page 2345, "Exact Confidence Limits for the Common Odds Ratio" on page 2363, and "Zelen's Exact Test for Equal Odds Ratios" on page 2363.

Definition of *p*-Values

For several tests in PROC FREQ, the test statistic is nonnegative, and large values of the test statistic indicate a departure from the null hypothesis. Such nondirectional tests include the Pearson chi-square, the likelihood-ratio chi-square, the Mantel-Haenszel chi-square, Fisher's exact test for tables larger than 2×2, McNemar's test, and the one-way chi-square goodness-of-fit test. The exact *p*-value for a nondirectional test is the sum of probabilities for those tables having a test statistic greater than or equal to the value of the observed test statistic.

There are other tests where it might be appropriate to test against either a one-sided or a two-sided alternative hypothesis. For example, when you test the null hypothesis that the true parameter value equals 0 ($T = 0$), the alternative of interest might be one-sided ($T \leq 0$, or $T \geq 0$) or two-sided ($T \neq 0$). Such tests include the Pearson correlation coefficient, Spearman correlation coefficient, Jonckheere-Terpstra test, Cochran-Armitage test for trend, simple kappa coefficient, and weighted kappa coefficient. For these tests, PROC FREQ displays the right-sided *p*-value when the observed value of the test statistic is greater than its expected value. The right-sided *p*-value is the sum of probabilities for those tables for which the test statistic is greater than or equal to the observed test statistic. Otherwise, when the observed test statistic is less than or equal to the expected value, PROC FREQ displays the left-sided *p*-value. The left-sided *p*-value is the sum of probabilities for those tables for which the test statistic is less than or equal to the one observed. The one-sided *p*-value P_1 can be expressed as

$$P_1 = \begin{cases} \text{Prob}(\text{ Test Statistic} \geq t\ ) & \text{if } t > \mathrm{E}_0(T) \\ \text{Prob}(\text{ Test Statistic} \leq t\ ) & \text{if } t \leq \mathrm{E}_0(T) \end{cases}$$

where t is the observed value of the test statistic and $\mathrm{E}_0(T)$ is the expected value of the test statistic under the null hypothesis. PROC FREQ computes the two-sided *p*-value as the sum of the one-sided *p*-value and the corresponding area in the opposite tail of the distribution of the statistic, equidistant from the expected

value. The two-sided p-value P_2 can be expressed as

$$P_2 = \text{Prob}\left(|\text{Test Statistic} - \text{E}_0(T)| \geq |t - \text{E}_0(T)| \right)$$

If you specify the POINT option in the EXACT statement, PROC FREQ also displays exact point probabilities for the test statistics. The exact point probability is the exact probability that the test statistic equals the observed value.

Computational Resources

PROC FREQ uses relatively fast and efficient algorithms for exact computations. These recently developed algorithms, together with improvements in computer power, now make it feasible to perform exact computations for data sets where previously only asymptotic methods could be applied. Nevertheless, there are still large problems that might require a prohibitive amount of time and memory for exact computations, depending on the speed and memory available on your computer. For large problems, consider whether exact methods are really needed or whether asymptotic methods might give results quite close to the exact results, while requiring much less computer time and memory. When asymptotic methods might not be sufficient for such large problems, consider using Monte Carlo estimation of exact p-values, as described in the section "Monte Carlo Estimation" on page 2369.

A formula does not exist that can predict in advance how much time and memory are needed to compute an exact p-value for a certain problem. The time and memory required depend on several factors, including which test is being performed, the total sample size, the number of rows and columns, and the specific arrangement of the observations into table cells. Generally, larger problems (in terms of total sample size, number of rows, and number of columns) tend to require more time and memory. Additionally, for a fixed total sample size, time and memory requirements tend to increase as the number of rows and columns increases, because this corresponds to an increase in the number of tables in the reference set. Also for a fixed sample size, time and memory requirements increase as the marginal row and column totals become more homogeneous. See Agresti, Mehta, and Patel (1990) and Gail and Mantel (1977) for more information.

At any time while PROC FREQ is computing exact p-values, you can terminate the computations by pressing the system interrupt key sequence (see the *SAS Companion* for your system) and choosing to stop computations. After you terminate exact computations, PROC FREQ completes all other remaining tasks. The procedure produces the requested output and reports missing values for any exact p-values that were not computed by the time of termination.

You can also use the MAXTIME= option in the EXACT statement to limit the amount of time PROC FREQ uses for exact computations. You specify a MAXTIME= value that is the maximum amount of clock time (in seconds) that PROC FREQ can use to compute an exact p-value. If PROC FREQ does not finish computing an exact p-value within that time, it terminates the computation and completes all other remaining tasks.

Monte Carlo Estimation

If you specify the option MC in the EXACT statement, PROC FREQ computes Monte Carlo estimates of the exact p-values instead of directly computing the exact p-values. Monte Carlo estimation can be useful for large problems that require a great amount of time and memory for exact computations but for which asymptotic approximations might not be sufficient. To describe the precision of each Monte Carlo estimate, PROC FREQ provides the asymptotic standard error and $100(1-\alpha)\%$ confidence limits. The confidence level α is determined by the ALPHA= option in the EXACT statement, which, by default, equals 0.01 and

produces 99% confidence limits. The N=*n* option in the EXACT statement specifies the number of samples that PROC FREQ uses for Monte Carlo estimation; the default is 10000 samples. You can specify a larger value for *n* to improve the precision of the Monte Carlo estimates. Because larger values of *n* generate more samples, the computation time increases. Alternatively, you can specify a smaller value of *n* to reduce the computation time.

To compute a Monte Carlo estimate of an exact *p*-value, PROC FREQ generates a random sample of tables with the same total sample size, row totals, and column totals as the observed table. PROC FREQ uses the algorithm of Agresti, Wackerly, and Boyett (1979), which generates tables in proportion to their hypergeometric probabilities conditional on the marginal frequencies. For each sample table, PROC FREQ computes the value of the test statistic and compares it to the value for the observed table. When estimating a right-sided *p*-value, PROC FREQ counts all sample tables for which the test statistic is greater than or equal to the observed test statistic. Then the *p*-value estimate equals the number of these tables divided by the total number of tables sampled.

$$\begin{aligned} \hat{P}_{MC} &= M \;/\; N \\ M &= \text{number of samples with (Test Statistic} \geq t) \\ N &= \text{total number of samples} \\ t &= \text{observed Test Statistic} \end{aligned}$$

PROC FREQ computes left-sided and two-sided *p*-value estimates in a similar manner. For left-sided *p*-values, PROC FREQ evaluates whether the test statistic for each sampled table is less than or equal to the observed test statistic. For two-sided *p*-values, PROC FREQ examines the sample test statistics according to the expression for P_2 given in the section "Definition of *p*-Values" on page 2368.

The variable M is a binomially distributed variable with N trials and success probability p. It follows that the asymptotic standard error of the Monte Carlo estimate is

$$\text{se}(\hat{P}_{MC}) = \sqrt{\hat{P}_{MC}\,(1 - \hat{P}_{MC}) \;/\; (N - 1)}$$

PROC FREQ constructs asymptotic confidence limits for the *p*-values according to

$$\hat{P}_{MC} \pm \left(z_{\alpha/2} \times \text{se}(\hat{P}_{MC}) \right)$$

where $z_{\alpha/2}$ is the $100(1 - \alpha/2)$th percentile of the standard normal distribution and the confidence level α is determined by the ALPHA= option in the EXACT statement.

When the Monte Carlo estimate $\hat{P}_{MC}$ equals 0, PROC FREQ computes the confidence limits for the *p*-value as

$$(0, \;\; 1 - \alpha^{(1/N)})$$

When the Monte Carlo estimate $\hat{P}_{MC}$ equals 1, PROC FREQ computes the confidence limits as

$$(\alpha^{(1/N)}, \;\; 1)$$

Computational Resources

For each variable in a table request, PROC FREQ stores all of the levels in memory. If all variables are numeric and not formatted, this requires about 84 bytes for each variable level. When there are character variables or formatted numeric variables, the memory that is required depends on the formatted variable lengths, with longer formatted lengths requiring more memory. The number of levels for each variable is limited only by the largest integer that your operating environment can store.

For any single crosstabulation table requested, PROC FREQ builds the entire table in memory, regardless of whether the table has zero cell counts. Thus, if the numeric variables A, B, and C each have 10 levels, PROC FREQ requires 2520 bytes to store the variable levels for the table request A*B*C, as follows:

```
3 variables * 10 levels/variable * 84 bytes/level
```

In addition, PROC FREQ requires 8000 bytes to store the table cell frequencies

```
1000 cells * 8 bytes/cell
```

even though there might be only 10 observations.

When the variables have many levels or when there are many multiway tables, your computer might not have enough memory to construct the tables. If PROC FREQ runs out of memory while constructing tables, it stops collecting levels for the variable with the most levels and returns the memory that is used by that variable. The procedure then builds the tables that do not contain the disabled variables.

If there is not enough memory for your table request and if increasing the available memory is impractical, you can reduce the number of multiway tables or variable levels. If you are not using the CMH or AGREE option in the TABLES statement to compute statistics across strata, reduce the number of multiway tables by using PROC SORT to sort the data set by one or more of the variables or by using the DATA step to create an index for the variables. Then remove the sorted or indexed variables from the TABLES statement and include a BY statement that uses these variables. You can also reduce memory requirements by using a FORMAT statement in the PROC FREQ step to reduce the number of levels. Additionally, reducing the formatted variable lengths reduces the amount of memory that is needed to store the variable levels. For more information about using formats, see the section "Grouping with Formats" on page 2309.

Output Data Sets

PROC FREQ produces two types of output data sets that you can use with other statistical and reporting procedures. You can request these data sets as follows:

- Specify the OUT= option in a TABLES statement. This creates an output data set that contains frequency or crosstabulation table counts and percentages
- Specify an OUTPUT statement. This creates an output data set that contains statistics.

PROC FREQ does not display the output data sets. Use PROC PRINT, PROC REPORT, or any other SAS reporting tool to display an output data set.

In addition to these two output data sets, you can create a SAS data set from any piece of PROC FREQ output by using the Output Delivery System. See the section "ODS Table Names" on page 2382 for more information.

Contents of the TABLES Statement Output Data Set

The OUT= option in the TABLES statement creates an output data set that contains one observation for each combination of variable values (or table cell) in the last table request. By default, each observation contains the frequency and percentage for the table cell. When the input data set contains missing values, the output data set also contains an observation with the frequency of missing values. The output data set includes the following variables:

- BY variables
- table request variables, such as A, B, C, and D in the table request A*B*C*D
- COUNT, which contains the table cell frequency
- PERCENT, which contains the table cell percentage

If you specify the OUTEXPECT option in the TABLES statement for a two-way or multiway table, the output data set also includes expected frequencies. If you specify the OUTPCT option for a two-way or multiway table, the output data set also includes row, column, and table percentages. The additional variables are as follows:

- EXPECTED, which contains the expected frequency
- PCT_TABL, which contains the percentage of two-way table frequency, for n-way tables where $n > 2$
- PCT_ROW, which contains the percentage of row frequency
- PCT_COL, which contains the percentage of column frequency

If you specify the OUTCUM option in the TABLES statement for a one-way table, the output data set also includes cumulative frequencies and cumulative percentages. The additional variables are as follows:

- CUM_FREQ, which contains the cumulative frequency
- CUM_PCT, which contains the cumulative percentage

The OUTCUM option has no effect for two-way or multiway tables.

The following PROC FREQ statements create an output data set of frequencies and percentages:

```
proc freq;
   tables A A*B / out=D;
run;
```

The output data set D contains frequencies and percentages for the table of A by B, which is the last table request listed in the TABLES statement. If A has two levels (1 and 2), B has three levels (1,2, and 3), and no table cell count is zero or missing, then the output data set D includes six observations, one for each combination of A and B levels. The first observation corresponds to A=1 and B=1; the second observation corresponds to A=1 and B=2; and so on. The data set includes the variables COUNT and PERCENT. The value of COUNT is the number of observations with the given combination of A and B levels. The value of PERCENT is the percentage of the total number of observations with that A and B combination.

When PROC FREQ combines different variable values into the same formatted level, the output data set contains the smallest internal value for the formatted level. For example, suppose a variable X has the values 1.1., 1.4, 1.7, 2.1, and 2.3. When you submit the statement

```
format X 1.;
```

in a PROC FREQ step, the formatted levels listed in the frequency table for X are 1 and 2. If you create an output data set with the frequency counts, the internal values of the levels of X are 1.1 and 1.7. To report the internal values of X when you display the output data set, use a format of 3.1 for X.

Contents of the OUTPUT Statement Output Data Set

The OUTPUT statement creates a SAS data set that contains the statistics that PROC FREQ computes for the last table request. You specify which statistics to store in the output data set. There is an observation with the specified statistics for each stratum or two-way table. If PROC FREQ computes summary statistics for a stratified table, the output data set also contains a summary observation with those statistics.

The OUTPUT data set can include the following variables.

- BY variables
- variables that identify the stratum, such as A and B in the table request A*B*C*D
- variables that contain the specified statistics

The output data set also includes variables with the *p*-values and degrees of freedom, asymptotic standard error (ASE), or confidence limits when PROC FREQ computes these values for a specified statistic.

The variable names for the specified statistics in the output data set are the names of the options enclosed in underscores. PROC FREQ forms variable names for the corresponding *p*-values, degrees of freedom, or confidence limits by combining the name of the option with the appropriate prefix from the following list:

DF_	degrees of freedom
E_	asymptotic standard error (ASE)
L_	lower confidence limit
U_	upper confidence limit
E0_	ASE under the null hypothesis
Z_	standardized value
P_	*p*-value
P2_	two-sided *p*-value
PL_	left-sided *p*-value
PR_	right-sided *p*-value

XP_	exact *p*-value
XP2_	exact two-sided *p*-value
XPL_	exact left-sided *p*-value
XPR_	exact right-sided *p*-value
XPT_	exact point probability
XL_	exact lower confidence limit
XU_	exact upper confidence limit

For example, variable names created for the Pearson chi-square, its degrees of freedom, and its *p*-values are _PCHI_, DF_PCHI, and P_PCHI, respectively.

If the length of the prefix plus the statistic option exceeds eight characters, PROC FREQ truncates the option so that the name of the new variable is eight characters long.

Displayed Output

Number of Variable Levels Table

If you specify the NLEVELS option in the PROC FREQ statement, PROC FREQ displays the "Number of Variable Levels" table. This table provides the number of levels for all variables named in the TABLES statements. PROC FREQ determines the variable levels from the formatted variable values. See "Grouping with Formats" on page 2309 for details. The "Number of Variable Levels" table contains the following information:

- Variable name
- Levels, which is the total number of levels of the variable
- Number of Nonmissing Levels, if there are missing levels for any of the variables
- Number of Missing Levels, if there are missing levels for any of the variables

One-Way Frequency Tables

PROC FREQ displays one-way frequency tables for all one-way table requests in the TABLES statements, unless you specify the NOPRINT option in the PROC statement or the NOPRINT option in the TABLES statement. For a one-way table showing the frequency distribution of a single variable, PROC FREQ displays the name of the variable and its values. For each variable value or level, PROC FREQ displays the following information:

- Frequency count, which is the number of observations in the level
- Test Frequency count, if you specify the CHISQ and TESTF= options to request a chi-square goodness-of-fit test for specified frequencies
- Percent, which is the percentage of the total number of observations. (The NOPERCENT option suppresses this information.)

- Test Percent, if you specify the CHISQ and TESTP= options to request a chi-square goodness-of-fit test for specified percents. (The NOPERCENT option suppresses this information.)
- Cumulative Frequency count, which is the sum of the frequency counts for that level and all other levels listed above it in the table. The last cumulative frequency is the total number of nonmissing observations. (The NOCUM option suppresses this information.)
- Cumulative Percent, which is the percentage of the total number of observations in that level and in all other levels listed above it in the table. (The NOCUM or the NOPERCENT option suppresses this information.)

The one-way table also displays the Frequency Missing, which is the number of observations with missing values.

Statistics for One-Way Frequency Tables

For one-way tables, two statistical options are available in the TABLES statement. The CHISQ option provides a chi-square goodness-of-fit test, and the BINOMIAL option provides binomial proportion statistics and tests. PROC FREQ displays the following information, unless you specify the NOPRINT option in the PROC FREQ statement:

- If you specify the CHISQ option for a one-way table, PROC FREQ provides a chi-square goodness-of-fit test, displaying the Chi-Square statistic, the degrees of freedom (DF), and the probability value (Pr > ChiSq). If you specify the CHISQ option in the EXACT statement, PROC FREQ also displays the exact probability value for this test. If you specify the POINT option with the CHISQ option in the EXACT statement, PROC FREQ displays the exact point probability for the test statistic.
- If you specify the BINOMIAL option for a one-way table, PROC FREQ displays the estimate of the binomial Proportion, which is the proportion of observations in the first class listed in the one-way table. PROC FREQ also displays the asymptotic standard error (ASE) and the asymptotic (Wald) and exact (Clopper-Pearson) confidence limits by default. For the binomial proportion test, PROC FREQ displays the asymptotic standard error under the null hypothesis (ASE Under H0), the standardized test statistic (Z), and the one-sided and two-sided probability values.

  If you specify the BINOMIAL option in the EXACT statement, PROC FREQ also displays the exact one-sided and two-sided probability values for this test. If you specify the POINT option with the BINOMIAL option in the EXACT statement, PROC FREQ displays the exact point probability for the test.
- If you request additional binomial confidence limits by specifying *binomial-options*, PROC FREQ provides a table that displays the lower and upper confidence limits for each type that you request. In addition to the Wald and exact (Clopper-Pearson) confidence limits, you can request Agresti-Coull, Jeffreys, and Wilson (score) confidence limits for the binomial proportion.
- If you request a binomial noninferiority or superiority test by specifying the NONINF or SUP *binomial-option*, PROC FREQ displays the following information: the binomial Proportion, the test ASE (under H0 or Sample), the test statistic Z, the probability value, the noninferiority or superiority limit, and the test confidence limits. If you specify the BINOMIAL option in the EXACT statement, PROC FREQ also provides the exact probability value for the test, and exact test confidence limits.

- If you request a binomial equivalence test by specifying the EQUIV *binomial-option*, PROC FREQ displays the binomial Proportion and the test ASE (under H0 or Sample). PROC FREQ displays two one-sided tests (TOST) for equivalence, which include test statistics (Z) and probability values for the Lower and Upper tests, together with the Overall probability value. PROC FREQ also displays the equivalence limits and the test-based confidence limits. If you specify the BINOMIAL option in the EXACT statement, PROC FREQ provides exact probability values for the TOST and exact test-based confidence limits.

Multiway Tables

PROC FREQ displays all multiway table requests in the TABLES statements, unless you specify the NOPRINT option in the PROC FREQ statement or the NOPRINT option in the TABLES statement.

For two-way to multiway crosstabulation tables, the values of the last variable in the table request form the table columns. The values of the next-to-last variable form the rows. Each level (or combination of levels) of the other variables forms one stratum.

There are three ways to display multiway tables in PROC FREQ. By default, PROC FREQ displays multiway tables as separate two-way crosstabulation tables for each stratum of the multiway table. Also by default, PROC FREQ displays these two-way crosstabulation tables in table cell format. Alternatively, if you specify the CROSSLIST option, PROC FREQ displays the two-way crosstabulation tables in ODS column format. If you specify the LIST option, PROC FREQ displays multiway tables in list format, which presents the entire multiway crosstabulation in a single table.

Crosstabulation Tables

By default, PROC FREQ displays two-way crosstabulation tables in table cell format. The row variable values are listed down the side of the table, the column variable values are listed across the top of the table, and each row and column variable level combination forms a table cell.

Each cell of a crosstabulation table can contain the following information:

- Frequency, which is the number of observations in the table cell. (The NOFREQ option suppresses this information.)
- Expected frequency under the hypothesis of independence, if you specify the EXPECTED option
- Deviation of the cell frequency from the expected value, if you specify the DEVIATION option
- Cell Chi-Square, which is the cell's contribution to the total chi-square statistic, if you specify the CELLCHI2 option
- Tot Pct, which is the cell's percentage of the total multiway table frequency, for n-way tables when $n > 2$, if you specify the TOTPCT option
- Percent, which is the cell's percentage of the total (two-way table) frequency. (The NOPERCENT option suppresses this information.)
- Row Pct, or the row percentage, which is the cell's percentage of the total frequency for its row. (The NOROW option suppresses this information.)

- Col Pct, or column percentage, which is the cell's percentage of the total frequency for its column. (The NOCOL option suppresses this information.)
- Cumulative Col%, or cumulative column percentage, if you specify the CUMCOL option

The table also displays the Frequency Missing, which is the number of observations with missing values.

CROSSLIST Tables

If you specify the CROSSLIST option, PROC FREQ displays two-way crosstabulation tables in ODS column format. The CROSSLIST column format is different from the default crosstabulation table cell format, but the CROSSLIST table provides the same information (frequencies, percentages, and other statistics) as the default crosstabulation table.

In the CROSSLIST table format, the rows of the display correspond to the crosstabulation table cells, and the columns of the display correspond to descriptive statistics such as frequencies and percentages. Each table cell is identified by the values of its TABLES row and column variable levels, with all column variable levels listed within each row variable level. The CROSSLIST table also provides row totals, column totals, and overall table totals.

For a crosstabulation table in CROSSLIST format, PROC FREQ displays the following information:

- the row variable name and values
- the column variable name and values
- Frequency, which is the number of observations in the table cell. (The NOFREQ option suppresses this information.)
- Expected cell frequency under the hypothesis of independence, if you specify the EXPECTED option
- Deviation of the cell frequency from the expected value, if you specify the DEVIATION option
- Cell Chi-Square, which is the cell's contribution to the total chi-square statistic, if you specify the CELLCHI2 option
- Total Percent, which is the cell's percentage of the total multiway table frequency, for n-way tables when $n > 2$, if you specify the TOTPCT option
- Percent, which is the cell's percentage of the total (two-way table) frequency. (The NOPERCENT option suppresses this information.)
- Row Percent, which is the cell's percentage of the total frequency for its row. (The NOROW option suppresses this information.)
- Column Percent, the cell's percentage of the total frequency for its column. (The NOCOL option suppresses this information.)

The table also displays the Frequency Missing, which is the number of observations with missing values.

LIST Tables

If you specify the LIST option in the TABLES statement, PROC FREQ displays multiway tables in a list format rather than as crosstabulation tables. The LIST option displays the entire multiway table in one table, instead of displaying a separate two-way table for each stratum. The LIST option is not available when you also request statistical options. Unlike the default crosstabulation output, the LIST output does not display row percentages, column percentages, and optional information such as expected frequencies and cell chi-squares.

For a multiway table in list format, PROC FREQ displays the following information:

- the variable names and values
- Frequency, which is the number of observations in the level (with the indicated variable values)
- Percent, which is the level's percentage of the total number of observations. (The NOPERCENT option suppresses this information.)
- Cumulative Frequency, which is the accumulated frequency of the level and all other levels listed above it in the table. The last cumulative frequency in the table is the total number of nonmissing observations. (The NOCUM option suppresses this information.)
- Cumulative Percent, which is the accumulated percentage of the level and all other levels listed above it in the table. (The NOCUM or the NOPERCENT option suppresses this information.)

The table also displays the Frequency Missing, which is the number of observations with missing values.

Statistics for Multiway Tables

PROC FREQ computes statistical tests and measures for crosstabulation tables, depending on which statements and options you specify. You can suppress the display of these results by specifying the NOPRINT option in the PROC FREQ statement. With any of the following information, PROC FREQ also displays the Sample Size and the Frequency Missing.

- If you specify the SCOROUT option in the TABLES statement, PROC FREQ displays the Row Scores and Column Scores that it uses for statistical computations. The Row Scores table displays the row variable values and the Score corresponding to each value. The Column Scores table displays the column variable values and the corresponding Scores. PROC FREQ also identifies the score type used to compute the row and column scores. You can specify the score type with the SCORES= option in the TABLES statement.
- If you specify the CHISQ option, PROC FREQ displays the following statistics for each two-way table: Pearson Chi-Square, Likelihood-Ratio Chi-Square, Continuity-Adjusted Chi-Square (for 2×2 tables), Mantel-Haenszel Chi-Square, the Phi Coefficient, the Contingency Coefficient, and Cramer's V. For each test statistic, PROC FREQ also displays the degrees of freedom (DF) and the probability value (Prob).
- If you specify the CHISQ option for 2×2 tables, PROC FREQ also displays Fisher's exact test. The test output includes the cell (1,1) frequency (F), the exact left-sided and right-sided probability values, the table probability (P), and the exact two-sided probability value.

- If you specify the FISHER option in the TABLES statement (or, equivalently, the FISHER option in the EXACT statement), PROC FREQ displays Fisher's exact test for tables larger than 2×2. The test output includes the table probability (P) and the probability value. In addition, PROC FREQ displays the CHISQ output listed earlier, even if you do not also specify the CHISQ option.

- If you specify the PCHI, LRCHI, or MHCHI option in the EXACT statement, PROC FREQ displays the corresponding exact test: Pearson Chi-Square, Likelihood-Ratio Chi-Square, or Mantel-Haenszel Chi-Square, respectively. The test output includes the test statistic, the degrees of freedom (DF), and the asymptotic and exact probability values. If you also specify the POINT option in the EXACT statement, PROC FREQ displays the point probability for each exact test requested. If you specify the CHISQ option in the EXACT statement, PROC FREQ displays exact probability values for all three of these chi-square tests.

- If you specify the MEASURES option, PROC FREQ displays the following statistics and their asymptotic standard errors (ASE) for each two-way table: Gamma, Kendall's Tau-*b*, Stuart's Tau-*c*, Somers' $D(C|R)$, Somers' $D(R|C)$, Pearson Correlation, Spearman Correlation, Lambda Asymmetric $(C|R)$, Lambda Asymmetric $(R|C)$, Lambda Symmetric, Uncertainty Coefficient $(C|R)$, Uncertainty Coefficient $(R|C)$, and Uncertainty Coefficient Symmetric. If you specify the CL option, PROC FREQ also displays confidence limits for these measures.

- If you specify the PLCORR option, PROC FREQ displays the tetrachoric correlation for 2×2 tables or the polychoric correlation for larger tables. In addition, PROC FREQ displays the MEASURES output listed earlier, even if you do not also specify the MEASURES option.

- If you specify the GAMMA, KENTB, STUTC, SMDCR, SMDRC, PCORR, or SCORR option in the TEST statement, PROC FREQ displays asymptotic tests for Gamma, Kendall's Tau-*b*, Stuart's Tau-*c*, Somers' $D(C|R)$, Somers' $D(R|C)$, the Pearson Correlation, or the Spearman Correlation, respectively. If you specify the MEASURES option in the TEST statement, PROC FREQ displays all these asymptotic tests. The test output includes the statistic, its asymptotic standard error (ASE), Confidence Limits, the ASE under the null hypothesis H0, the standardized test statistic (Z), and the one-sided and two-sided probability values.

- If you specify the KENTB, STUTC, SMDCR, SMDRC, PCORR, or SCORR option in the EXACT statement, PROC FREQ displays asymptotic and exact tests for the corresponding measure of association: Kendall's Tau-*b*, Stuart's Tau-*c*, Somers' $D(C|R)$, Somers' $D(R|C)$, the Pearson Correlation, or the Spearman correlation, respectively. The test output includes the correlation, its asymptotic standard error (ASE), Confidence Limits, the ASE under the null hypothesis H0, the standardized test statistic (Z), and the asymptotic and exact one-sided and two-sided probability values. If you also specify the POINT option in the EXACT statement, PROC FREQ displays the point probability for each exact test requested.

- If you specify the RISKDIFF option for 2×2 tables, PROC FREQ displays the Column 1 and Column 2 Risk Estimates. For each column, PROC FREQ displays the Row 1 Risk, Row 2 Risk, Total Risk, and Risk Difference, together with their asymptotic standard errors (ASE) and Asymptotic Confidence Limits. PROC FREQ also displays Exact Confidence Limits for the Row 1 Risk, Row 2 Risk, and Total Risk. If you specify the RISKDIFF option in the EXACT statement, PROC FREQ provides unconditional Exact Confidence Limits for the Risk Difference.

- If you specify the RISKDIFF(CL=) option for 2×2 tables, PROC FREQ displays the Proportion Difference Confidence Limits. For each confidence limit Type that you request (Exact, Farrington-

Manning, Hauck-Anderson, Newcombe Score, or Wald), PROC FREQ displays the Lower and Upper Confidence Limits.

- If you request a noninferiority or superiority test for the proportion difference (RISKDIFF) by specifying the NONINF or SUP *riskdiff-option*, and if you specify METHOD=HA (Hauck-Anderson), METHOD=FM (Farrington-Manning), or METHOD=WALD (Wald), PROC FREQ displays the following information: the Proportion Difference, the test ASE (H0, Sample, Sample H-A, or FM, depending on the method you specify), the test statistic Z, the probability value, the Noninferiority or Superiority Limit, and the test-based Confidence Limits. If you specify METHOD=NEWCOMBE (Newcombe score), PROC FREQ displays the Proportion Difference, the Noninferiority or Superiority Limit, and the Newcombe Confidence Limits.

- If you request an equivalence test for the proportion difference (RISKDIFF) by specifying the EQUIV *riskdiff-option*, and if you specify METHOD=HA (Hauck-Anderson), METHOD=FM (Farrington-Manning), or METHOD=WALD (Wald), PROC FREQ displays the following information: the Proportion Difference and the test ASE (H0, Sample, Sample H-A, or FM, depending on the method you specify). PROC FREQ displays a two one-sided test (TOST) for equivalence, which includes test statistics (Z) and probability values for the Lower and Upper tests, together with the Overall probability value. PROC FREQ also displays the Equivalence Limits and the test-based Confidence Limits. If you specify METHOD=NEWCOMBE (Newcombe score), PROC FREQ displays the Proportion Difference, the Equivalence Limits, and the score Confidence Limits.

- If you request an equality test for the proportion difference (RISKDIFF) by specifying the EQUAL *riskdiff-option*, PROC FREQ displays the following information: the Proportion Difference and the test ASE (H0 or Sample), the test statistic Z, the One-Sided probability value (Pr > Z or Pr < Z), and the Two-Sided probability value, Pr > |Z|.

- If you specify the MEASURES option or the RELRISK option for 2×2 tables, PROC FREQ displays Estimates of the Relative Risk for Case-Control and Cohort studies, together with their Confidence Limits. These measures are also known as the Odds Ratio and the Column 1 and 2 Relative Risks. If you specify the OR option in the EXACT statement, PROC FREQ also displays Exact Confidence Limits for the Odds Ratio. If you specify the RELRISK option in the EXACT statement, PROC FREQ displays unconditional Exact Confidence Limits for the Relative Risk.

- If you specify the TREND option, PROC FREQ displays the Cochran-Armitage Trend Test for tables that are $2 \times C$ or $R \times 2$. For this test, PROC FREQ gives the Statistic (Z) and the one-sided and two-sided probability values. If you specify the TREND option in the EXACT statement, PROC FREQ also displays the exact one-sided and two-sided probability values for this test. If you specify the POINT option with the TREND option in the EXACT statement, PROC FREQ displays the exact point probability for the test statistic.

- If you specify the JT option, PROC FREQ displays the Jonckheere-Terpstra Test, showing the Statistic (JT), the standardized test statistic (Z), and the one-sided and two-sided probability values. If you specify the JT option in the EXACT statement, PROC FREQ also displays the exact one-sided and two-sided probability values for this test. If you specify the POINT option with the JT option in the EXACT statement, PROC FREQ displays the exact point probability for the test statistic.

- If you specify the AGREE option and the PRINTKWT option, PROC FREQ displays the Kappa Coefficient Weights for square tables larger than 2×2.

- If you specify the AGREE option, for two-way tables PROC FREQ displays McNemar's Test and the Simple Kappa Coefficient for 2×2 tables. For square tables larger than 2×2, PROC FREQ displays Bowker's Test of Symmetry, the Simple Kappa Coefficient, and the Weighted Kappa Coefficient. For McNemar's Test and Bowker's Test of Symmetry, PROC FREQ displays the Statistic (S), the degrees of freedom (DF), and the probability value (Pr > S). If you specify the MCNEM option in the EXACT statement, PROC FREQ also displays the exact probability value for McNemar's test. If you specify the POINT option with the MCNEM option in the EXACT statement, PROC FREQ displays the exact point probability for the test statistic. For the simple and weighted kappa coefficients, PROC FREQ displays the kappa values, asymptotic standard errors (ASE), and Confidence Limits.

- If you specify the KAPPA or WTKAP option in the TEST statement, PROC FREQ displays asymptotic tests for the simple kappa coefficient or the weighted kappa coefficient, respectively. If you specify the AGREE option in the TEST statement, PROC FREQ displays both these asymptotic tests. The test output includes the kappa coefficient, its asymptotic standard error (ASE), Confidence Limits, the ASE under the null hypothesis H0, the standardized test statistic (Z), and the one-sided and two-sided probability values.

- If you specify the KAPPA or WTKAP option in the EXACT statement, PROC FREQ displays asymptotic and exact tests for the simple kappa coefficient or the weighted kappa coefficient, respectively. The test output includes the kappa coefficient, its asymptotic standard error (ASE), Confidence Limits, the ASE under the null hypothesis H0, the standardized test statistic (Z), and the asymptotic and exact one-sided and two-sided probability values. If you specify the POINT option in the EXACT statement, PROC FREQ displays the point probability for each exact test requested.

- If you specify the MC option in the EXACT statement, PROC FREQ displays Monte Carlo estimates for all exact p-values requested by *statistic-options* in the EXACT statement. The Monte Carlo output includes the p-value Estimate, its Confidence Limits, the Number of Samples used to compute the Monte Carlo estimate, and the Initial Seed for random number generation.

- If you specify the AGREE option, for multiple strata PROC FREQ displays Overall Simple and Weighted Kappa Coefficients, with their asymptotic standard errors (ASE) and Confidence Limits. PROC FREQ also displays Tests for Equal Kappa Coefficients, giving the Chi-Squares, degrees of freedom (DF), and probability values (Pr > ChiSq) for the Simple Kappa and Weighted Kappa. For multiple strata of 2×2 tables, PROC FREQ displays Cochran's Q, giving the Statistic (Q), the degrees of freedom (DF), and the probability value (Pr > Q).

- If you specify the CMH option, PROC FREQ displays Cochran-Mantel-Haenszel Statistics for the following three alternative hypotheses: Nonzero Correlation, Row Mean Scores Differ (ANOVA Statistic), and General Association. For each of these statistics, PROC FREQ gives the degrees of freedom (DF) and the probability value (Prob). If you specify the MANTELFLEISS option, PROC FREQ displays the Mantel-Fleiss Criterion for 2×2 tables. For 2×2 tables, PROC FREQ also displays Estimates of the Common Relative Risk for Case-Control and Cohort studies, together with their confidence limits. These include both Mantel-Haenszel and Logit stratum-adjusted estimates of the common Odds Ratio, Column 1 Relative Risk, and Column 2 Relative Risk. Also for 2×2 tables, PROC FREQ displays the Breslow-Day Test for Homogeneity of the Odds Ratios. For this test, PROC FREQ gives the Chi-Square, the degrees of freedom (DF), and the probability value (Pr > ChiSq).

- If you specify the CMH option in the TABLES statement and also specify the COMOR option in the EXACT statement, PROC FREQ displays exact confidence limits for the Common

Odds Ratio for multiple strata of 2 × 2 tables. PROC FREQ also displays the Exact Test of H0: Common Odds Ratio = 1. The test output includes the Cell (1,1) Sum (S), Mean of S Under H0, One-sided Pr <= S, and Point Pr = S. PROC FREQ also provides exact two-sided probability values for the test, computed according to the following three methods: 2 * One-sided, Sum of probabilities <= Point probability, and Pr >= |S - Mean|.

- If you specify the CMH option in the TABLES statement and also specify the EQOR option in the EXACT statement, PROC FREQ computes Zelen's exact test for equal odds ratios for $h \times 2 \times 2$ tables. PROC FREQ displays Zelen's test along with the asymptotic Breslow-Day test produced by the CMH option. PROC FREQ displays the test statistic, Zelen's Exact Test (P), and the probability value, Exact Pr <= P.
- If you specify the GAILSIMON option in the TABLES statement for a multiway 2 × 2 tables, PROC FREQ displays the Gail-Simon test for qualitative interactions. The display include the following statistics and their p-values: Q+ (Positive Risk Differences), Q- (Negative Risk Differences), and Q (Two-Sided).

ODS Table Names

PROC FREQ assigns a name to each table that it creates. You can use these names to refer to tables when you use the Output Delivery System (ODS) to select tables and create output data sets. For more information about ODS, see Chapter 20, "Using the Output Delivery System."

Table 36.14 lists the ODS table names together with their descriptions and the options required to produce the tables. Note that the ALL option in the TABLES statement invokes the CHISQ, MEASURES, and CMH options.

Table 36.14 ODS Tables Produced by PROC FREQ

ODS Table Name	Description	Statement	Option
BinomialCLs	Binomial confidence limits	TABLES	BINOMIAL(AC | J | W)
BinomialEquiv	Binomial equivalence analysis	TABLES	BINOMIAL(EQUIV)
BinomialEquivLimits	Binomial equivalence limits	TABLES	BINOMIAL(EQUIV)
BinomialEquivTest	Binomial equivalence test	TABLES	BINOMIAL(EQUIV)
BinomialNoninf	Binomial noninferiority test	TABLES	BINOMIAL(NONINF)
BinomialProp	Binomial proportion	TABLES	BINOMIAL
BinomialPropTest	Binomial proportion test	TABLES	BINOMIAL
BinomialSup	Binomial superiority test	TABLES	BINOMIAL(SUP)
BreslowDayTest	Breslow-Day test	TABLES	CMH ($h \times 2 \times 2$ table)
CMH	Cochran-Mantel-Haenszel statistics	TABLES	CMH
ChiSq	Chi-square tests	TABLES	CHISQ
CochransQ	Cochran's Q	TABLES	AGREE ($h \times 2 \times 2$ table)
ColScores	Column scores	TABLES	SCOROUT
CommonOddsRatioCl	Exact confidence limits for the common odds ratio	EXACT	COMOR ($h \times 2 \times 2$ table)

Table 36.14 *continued*

ODS Table Name	Description	Statement	Option
CommonOdds-RatioTest	Common odds ratio exact test	EXACT	COMOR ($h \times 2 \times 2$ table)
CommonRelRisks	Common relative risks	TABLES	CMH ($h \times 2 \times 2$ table)
CrossList	Crosstabulation table in column format	TABLES	CROSSLIST (n-way table, $n > 1$)
CrossTabFreqs	Crosstabulation table	TABLES	(n-way table, $n > 1$)
EqualKappaTest	Test for equal simple kappas	TABLES	AGREE ($h \times 2 \times 2$ table)
EqualKappaTests	Tests for equal kappas	TABLES	AGREE ($h \times r \times r$, $r > 2$)
EqualOddsRatios	Tests for equal odds ratios	EXACT	EQOR ($h \times 2 \times 2$ table)
GailSimon	Gail-Simon test	TABLES	GAILSIMON ($h \times 2 \times 2$ table)
FishersExact	Fisher's exact test	EXACT or TABLES or TABLES	FISHER FISHER or EXACT CHISQ (2×2 table)
FishersExactMC	Monte Carlo estimates for Fisher's exact test	EXACT	FISHER / MC
Gamma	Gamma	TEST	GAMMA
GammaTest	Gamma test	TEST	GAMMA
JTTest	Jonckheere-Terpstra test	TABLES	JT
JTTestMC	Monte Carlo estimates for Jonckheere-Terpstra exact test	EXACT	JT / MC
KappaStatistics	Kappa statistics	TABLES	AGREE, no TEST or EXACT ($r \times r$ table, $r > 2$)
KappaWeights	Kappa weights	TABLES	AGREE and PRINTKWT
List	List format multiway table	TABLES	LIST
LRChiSq	Likelihood-ratio chi-square exact test	EXACT	LRCHI
LRChiSqMC	Monte Carlo exact test for likelihood-ratio chi-square	EXACT	LRCHI / MC
MantelFleiss	Mantel-Fleiss criterion	TABLES	CMH(MF) ($h \times 2 \times 2$ table)
McNemarsTest	McNemar's test	TABLES	AGREE (2×2 table)
Measures	Measures of association	TABLES	MEASURES

Table 36.14 *continued*

ODS Table Name	Description	Statement	Option
MHChiSq	Mantel-Haenszel chi-square exact test	EXACT	MHCHI
MHChiSqMC	Monte Carlo exact test for Mantel-Haenszel chi-square	EXACT	MHCHI / MC
NLevels	Number of variable levels	PROC	NLEVELS
OddsRatioCL	Exact confidence limits for the odds ratio	EXACT	OR (2×2 table)
OneWayChiSq	One-way chi-square test	TABLES	CHISQ (one-way table)
OneWayChiSqMC	Monte Carlo exact test for one-way chi-square	EXACT	CHISQ / MC (one-way table)
OneWayFreqs	One-way frequencies	PROC or TABLES	(no TABLES stmt) (one-way table)
OverallKappa	Overall simple kappa	TABLES	AGREE ($h \times 2 \times 2$ table)
OverallKappas	Overall kappa coefficients	TABLES	AGREE ($h \times r \times r, r > 2$)
PdiffCLs	Proportion difference confidence limits	TABLES	RISKDIFF(CL=) (2×2 table)
PdiffEquiv	Equivalence analysis for the proportion difference	TABLES	RISKDIFF(EQUIV) (2×2 table)
PdiffEquivLimits	Equivalence limits for the proportion difference	TABLES	RISKDIFF(EQUIV) (2×2 table)
PdiffEquivTest	Equivalence test for the proportion difference	TABLES	RISKDIFF(EQUIV) (2×2 table)
PdiffNoninf	Noninferiority test for the proportion difference	TABLES	RISKDIFF(NONINF) (2×2 table)
PdiffSup	Superiority test for the proportion difference	TABLES	RISKDIFF(SUP) (2×2 table)
PdiffTest	Proportion difference test	TABLES	RISKDIFF(EQUAL) (2×2 table)
PearsonChiSq	Pearson chi-square exact test	EXACT	PCHI
PearsonChiSqMC	Monte Carlo exact test for Pearson chi-square	EXACT	PCHI / MC
PearsonCorr	Pearson correlation	TEST or EXACT	PCORR PCORR
PearsonCorrMC	Monte Carlo exact test for Pearson correlation	EXACT	PCORR / MC
PearsonCorrTest	Pearson correlation test	TEST or EXACT	PCORR PCORR
RelativeRisks	Relative risk estimates	TABLES	RELRISK or MEASURES (2×2 table)

Table 36.14 *continued*

ODS Table Name	Description	Statement	Option
RelRisk1CL	Exact confidence limits for column 1 relative risk	EXACT	RELRISK (2×2 table)
RelRisk2CL	Exact confidence limits for column 2 relative risk	EXACT	RELRISK (2×2 table)
RiskDiffCol1	Column 1 risk estimates	TABLES	RISKDIFF (2×2 table)
RiskDiffCol2	Column 2 risk estimates	TABLES	RISKDIFF (2×2 table)
RowScores	Row scores	TABLES	SCOROUT
SimpleKappa	Simple kappa coefficient	TEST	KAPPA
		or EXACT	KAPPA
SimpleKappaMC	Monte Carlo exact test for simple kappa	EXACT	KAPPA / MC
SimpleKappaTest	Simple kappa test	TEST	KAPPA
		or EXACT	KAPPA
SomersDCR	Somers' $D(C\|R)$	TEST	SMDCR
SomersDCRTest	Somers' $D(C\|R)$ test	TEST	SMDCR
SomersDRC	Somers' $D(R\|C)$	TEST	SMDRC
SomersDRCTest	Somers' $D(R\|C)$ test	TEST	SMDRC
SpearmanCorr	Spearman correlation	TEST	SCORR
		or EXACT	SCORR
SpearmanCorrMC	Monte Carlo exact test for Spearman correlation	EXACT	SCORR / MC
SpearmanCorrTest	Spearman correlation test	TEST	SCORR
		or EXACT	SCORR
SymmetryTest	Test of symmetry	TABLES	AGREE
TauB	Kendall's tau-b	TEST	KENTB
TauBTest	Kendall's tau-b test	TEST	KENTB
TauC	Stuart's tau-c	TEST	STUTC
TauCTest	Stuart's tau-c test	TEST	STUTC
TrendTest	Cochran-Armitage trend test	TABLES	TREND
TrendTestMC	Monte Carlo exact test for trend	EXACT	TREND / MC
WeightedKappa	Weighted kappa	TEST	WTKAP
		or EXACT	WTKAP
WeightedKappaMC	Monte Carlo exact test for weighted kappa	EXACT	WTKAP / MC
WeightedKappaTest	Weighted kappa test	TEST	WTKAP
		or EXACT	WTKAP

ODS Graphics

Statistical procedures use ODS Graphics to create graphs as part of their output. ODS Graphics is described in detail in Chapter 21, "Statistical Graphics Using ODS."

Before you create graphs, ODS Graphics must be enabled (for example, with the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 609 in Chapter 21, "Statistical Graphics Using ODS."

The overall appearance of graphs is controlled by ODS styles. Styles and other aspects of using ODS Graphics are discussed in the section "A Primer on ODS Statistical Graphics" on page 608 in Chapter 21, "Statistical Graphics Using ODS."

When ODS Graphics is enabled, you can request specific plots with the PLOTS= option in the TABLES statement. To produce a frequency plot or cumulative frequency plot, you must specify the FREQPLOT or CUMFREQPLOT *plot-request*, respectively, in the PLOTS= option. By default, PROC FREQ produces all other plots that are associated with the analyses that you request in the TABLES statement. You can suppress the default plots and request specific plots by using the PLOTS(ONLY)= option. See the description of the PLOTS= option for details.

PROC FREQ assigns a name to each graph that it creates with ODS Graphics. You can use these names to refer to the graphs. Table 36.15 lists the names of the graphs that PROC FREQ generates together with their descriptions, their PLOTS= options (*plot-requests*), and the TABLES statement options that are required to produce the graphs.

Table 36.15 Graphs Produced by PROC FREQ

ODS Graph Name	Description	PLOTS= Option	TABLES Statement Option
AgreePlot	Agreement plot	AGREEPLOT	AGREE ($r \times r$ table)
CumFreqPlot	Cumulative frequency plot	CUMFREQPLOT	One-way table request
DeviationPlot	Deviation plot	DEVIATIONPLOT	CHISQ (one-way table)
FreqPlot	Frequency plot	FREQPLOT	Any table request
KappaPlot	Kappa plot	KAPPAPLOT	AGREE ($h \times r \times r$ table)
ORPlot	Odds ratio plot	ODDSRATIOPLOT	MEASURES or RELRISK ($h \times 2 \times 2$ table)
RelRiskPlot	Relative risk plot	RELRISKPLOT	MEASURES or RELRISK ($h \times 2 \times 2$ table)
RiskDiffPlot	Risk difference plot	RISKDIFFPLOT	RISKDIFF ($h \times 2 \times 2$ table)
WtKappaPlot	Weighted kappa plot	WTKAPPAPLOT	AGREE ($h \times r \times r$ table, $r > 2$)

Examples: FREQ Procedure

Example 36.1: Output Data Set of Frequencies

The eye and hair color of children from two different regions of Europe are recorded in the data set Color. Instead of recording one observation per child, the data are recorded as cell counts, where the variable Count contains the number of children exhibiting each of the 15 eye and hair color combinations. The data set does not include missing combinations.

The following DATA step statements create the SAS data set Color:

```
data Color;
   input Region Eyes $ Hair $ Count @@;
   label Eyes  ='Eye Color'
         Hair  ='Hair Color'
         Region='Geographic Region';
   datalines;
1 blue  fair   23  1 blue  red     7  1 blue  medium 24
1 blue  dark   11  1 green fair   19  1 green red     7
1 green medium 18  1 green dark   14  1 brown fair   34
1 brown red     5  1 brown medium 41  1 brown dark   40
1 brown black   3  2 blue  fair   46  2 blue  red    21
2 blue  medium 44  2 blue  dark   40  2 blue  black   6
2 green fair   50  2 green red    31  2 green medium 37
2 green dark   23  2 brown fair   56  2 brown red    42
2 brown medium 53  2 brown dark   54  2 brown black  13
;
```

The following PROC FREQ statements read the Color data set and create an output data set that contains the frequencies, percentages, and expected cell frequencies of the two-way table of Eyes by Hair. The TABLES statement requests three tables: a frequency table for Eyes, a frequency table for Hair, and a crosstabulation table for Eyes by Hair. The OUT= option creates the FreqCount data set, which contains the crosstabulation table frequencies. The OUTEXPECT option outputs the expected table cell frequencies to FreqCount, and the SPARSE option includes zero cell frequencies in the output data set. The WEIGHT statement specifies that the variable Count contains the observation weights. These statements create Output 36.1.1 through Output 36.1.3.

```
proc freq data=Color;
   tables Eyes Hair Eyes*Hair / out=FreqCount outexpect sparse;
   weight Count;
   title 'Eye and Hair Color of European Children';
run;

proc print data=FreqCount noobs;
   title2 'Output Data Set from PROC FREQ';
run;
```

Output 36.1.1 displays the two frequency tables produced by PROC FREQ: one showing the distribution of eye color, and one showing the distribution of hair color. By default, PROC FREQ lists the variables values in alphabetical order. The 'Eyes*Hair' specification produces a crosstabulation table, shown in Output 36.1.2, with eye color defining the table rows and hair color defining the table columns. A zero cell frequency for green eyes and black hair indicates that this eye and hair color combination does not occur in the data.

The output data set FreqCount (Output 36.1.3) contains frequency counts and percentages for the last table requested in the TABLES statement, Eyes by Hair. Because the SPARSE option is specified, the data set includes the observation with a zero frequency. The variable Expected contains the expected frequencies, as requested by the OUTEXPECT option.

Output 36.1.1 Frequency Tables

Eye and Hair Color of European Children

The FREQ Procedure

Eye Color

Eyes	Frequency	Percent	Cumulative Frequency	Cumulative Percent
blue	222	29.13	222	29.13
brown	341	44.75	563	73.88
green	199	26.12	762	100.00

Hair Color

Hair	Frequency	Percent	Cumulative Frequency	Cumulative Percent
black	22	2.89	22	2.89
dark	182	23.88	204	26.77
fair	228	29.92	432	56.69
medium	217	28.48	649	85.17
red	113	14.83	762	100.00

Output 36.1.2 Crosstabulation Table

Table of Eyes by Hair

Eyes(Eye Color) Hair(Hair Color)

Frequency Percent Row Pct Col Pct	black	dark	fair	medium	red	Total
blue	6 0.79 2.70 27.27	51 6.69 22.97 28.02	69 9.06 31.08 30.26	68 8.92 30.63 31.34	28 3.67 12.61 24.78	222 29.13
brown	16 2.10 4.69 72.73	94 12.34 27.57 51.65	90 11.81 26.39 39.47	94 12.34 27.57 43.32	47 6.17 13.78 41.59	341 44.75
green	0 0.00 0.00 0.00	37 4.86 18.59 20.33	69 9.06 34.67 30.26	55 7.22 27.64 25.35	38 4.99 19.10 33.63	199 26.12
Total	22 2.89	182 23.88	228 29.92	217 28.48	113 14.83	762 100.00

Output 36.1.3 Output Data Set of Frequencies

Eye and Hair Color of European Children
Output Data Set from PROC FREQ

Eyes	Hair	COUNT	EXPECTED	PERCENT
blue	black	6	6.409	0.7874
blue	dark	51	53.024	6.6929
blue	fair	69	66.425	9.0551
blue	medium	68	63.220	8.9239
blue	red	28	32.921	3.6745
brown	black	16	9.845	2.0997
brown	dark	94	81.446	12.3360
brown	fair	90	102.031	11.8110
brown	medium	94	97.109	12.3360
brown	red	47	50.568	6.1680
green	black	0	5.745	0.0000
green	dark	37	47.530	4.8556
green	fair	69	59.543	9.0551
green	medium	55	56.671	7.2178
green	red	38	29.510	4.9869

Example 36.2: Frequency Dot Plots

This example produces frequency dot plots for the children's eye and hair color data from Example 36.1.

PROC FREQ produces plots by using ODS Graphics to create graphs as part of the procedure output. Frequency plots are available for any frequency or crosstabulation table request. You can display frequency plots as bar charts or dot plots. You can use *plot-options* to specify the orientation (vertical or horizontal), scale, and layout of the plots.

The following PROC FREQ statements request frequency tables and dot plots. The first TABLES statement requests a one-way frequency table of Hair and a crosstabulation table of Eyes by Hair. The PLOTS= option requests frequency plots for the tables, and the TYPE=DOTPLOT *plot-option* specifies dot plots. By default, frequency plots are produced as bar charts. ODS Graphics must be enabled before producing plots.

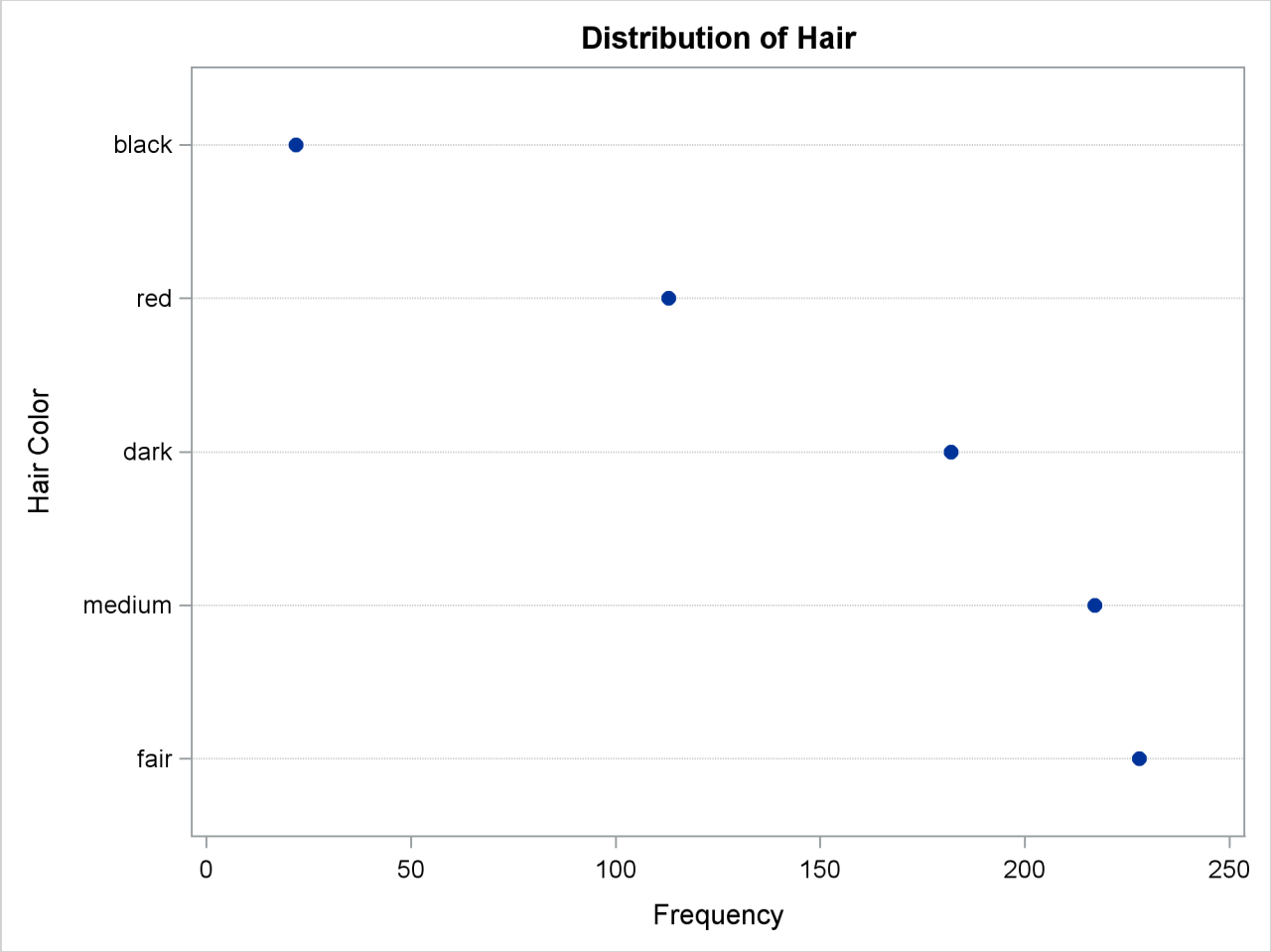
The second TABLES statement requests a crosstabulation table of Region by Hair and a frequency dot plot for this table. The SCALE=PERCENT *plot-option* plots percentages instead of frequency counts. SCALE=LOG and SCALE=SQRT *plot-options* are also available to plot log frequencies and square roots of frequencies, respectively.

The ORDER=FREQ option in the PROC FREQ statement orders the variable levels by frequency. This order applies to the frequency and crosstabulation table displays and also to the corresponding frequency plots.

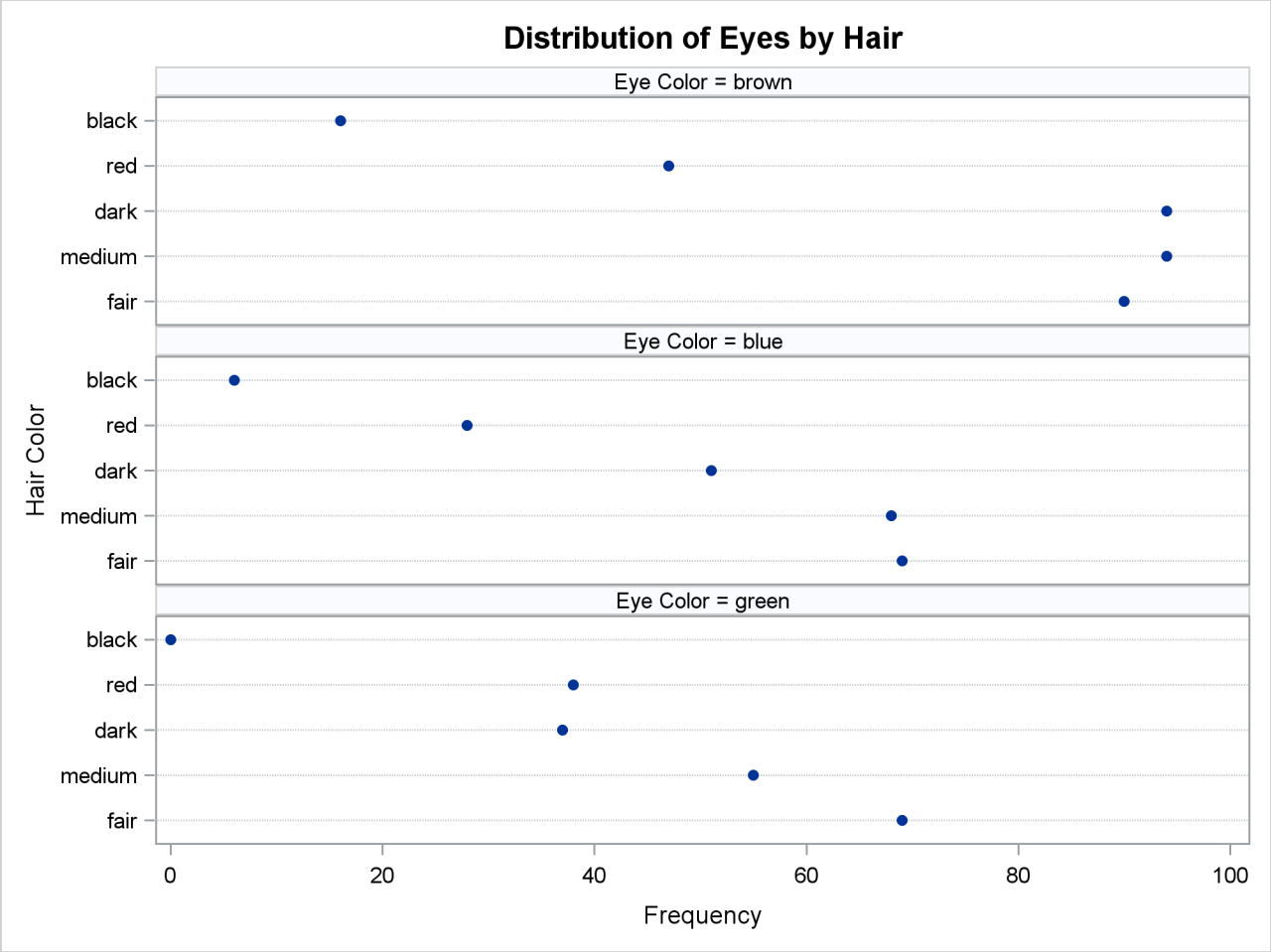
```
ods graphics on;
proc freq data=Color order=freq;
   tables Hair Eyes*Hair / plots=freqplot(type=dotplot);
   tables Region*Hair / plots=freqplot(type=dotplot scale=percent);
   weight Count;
   title 'Eye and Hair Color of European Children';
run;
ods graphics off;
```

Output 36.2.1, Output 36.2.2, and Output 36.2.3 display the dot plots produced by PROC FREQ. By default, the orientation of dot plots is horizontal, which places the variable levels on the Y axis. You can specify the ORIENT=VERTICAL *plot-option* to request a vertical orientation. For two-way plots, you can use the TWOWAY= *plot-option* to specify the plot layout. The default layout (shown in Output 36.2.2 and Output 36.2.3) is GROUPVERTICAL. Two-way layouts STACKED and GROUPHORIZONTAL are also available.

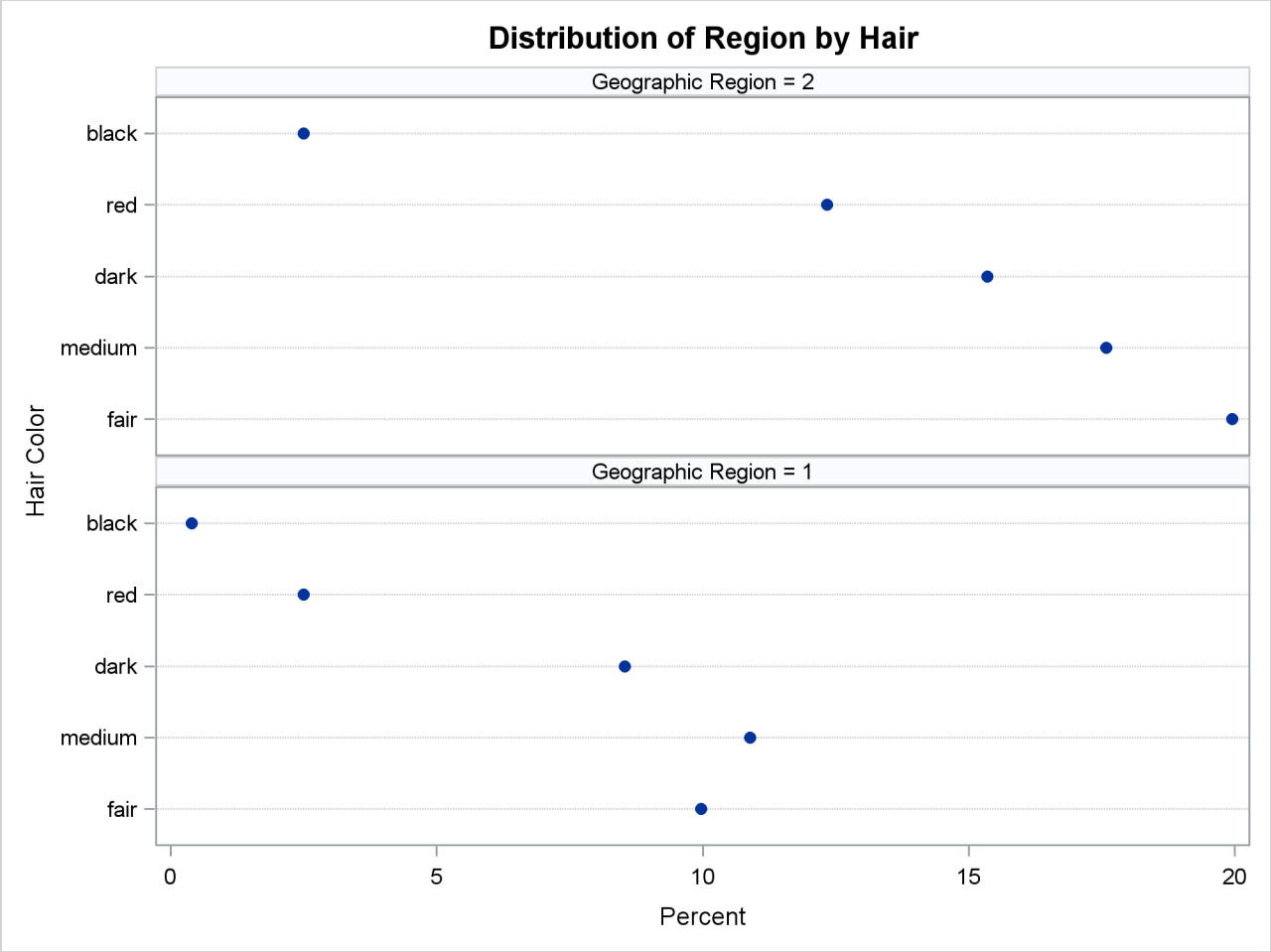
Output 36.2.1 One-Way Frequency Dot Plot



Output 36.2.2 Two-Way Frequency Dot Plot



Output 36.2.3 Two-Way Percent Dot Plot



Example 36.3: Chi-Square Goodness-of-Fit Tests

This example examines whether the children's hair color (from Example 36.1) has a specified multinomial distribution for the two geographical regions. The hypothesized distribution of hair color is 30% fair, 12% red, 30% medium, 25% dark, and 3% black.

In order to test the hypothesis for each region, the data are first sorted by Region. Then the FREQ procedure uses a BY statement to produce a separate table for each BY group (Region). The option ORDER=DATA orders the variable values (hair color) in the frequency table by their order in the input data set. The TABLES statement requests a frequency table for hair color, and the option NOCUM suppresses the display of the cumulative frequencies and percentages.

The CHISQ option requests a chi-square goodness-of-fit test for the frequency table of Hair. The TESTP= option specifies the hypothesized (or test) percentages for the chi-square test; the number of percentages listed equals the number of table levels, and the percentages sum to 100%. The TESTP= percentages are listed in the same order as the corresponding variable levels appear in frequency table.

The PLOTS= option requests a deviation plot, which is associated with the CHISQ option and displays the relative deviations from the test frequencies. The TYPE=DOTPLOT *plot-option* requests a dot plot instead of the default type, which is a bar chart. ODS Graphics must be enabled before producing plots. These statements produce Output 36.3.1 through Output 36.3.4.

```
proc sort data=Color;
   by Region;
run;

ods graphics on;
proc freq data=Color order=data;
   tables Hair / nocum chisq testp=(30 12 30 25 3)
                 plots(only)=deviationplot(type=dotplot);
   weight Count;
   by Region;
   title 'Hair Color of European Children';
run;
ods graphics off;
```

Output 36.3.1 Frequency Table and Chi-Square Test for Region 1

Hair Color of European Children

Geographic Region=1

The FREQ Procedure

Hair Color

Hair	Frequency	Percent	Test Percent
fair	76	30.89	30.00
red	19	7.72	12.00
medium	83	33.74	30.00
dark	65	26.42	25.00
black	3	1.22	3.00

Geographic Region=1

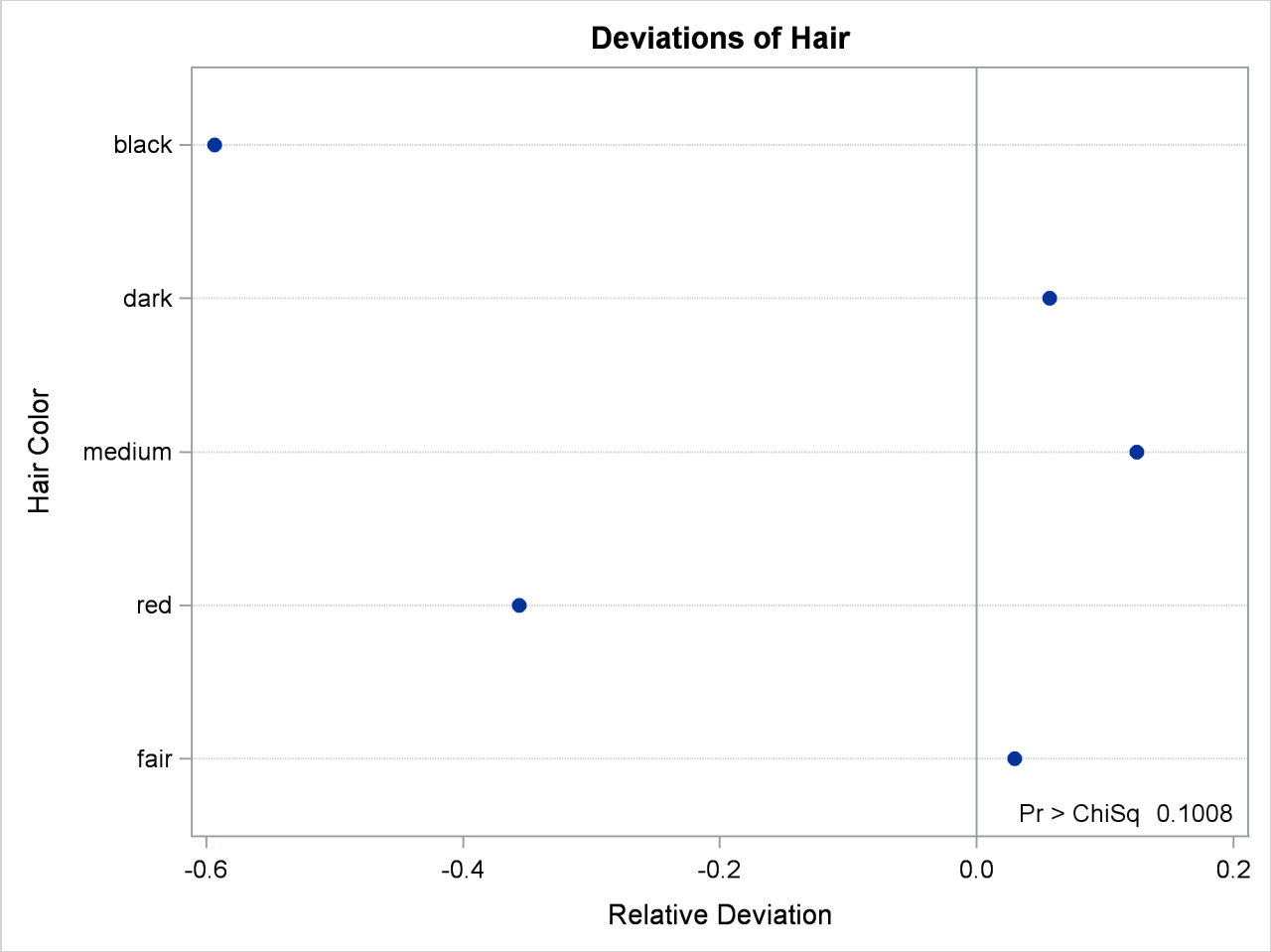
Chi-Square Test for Specified Proportions

Chi-Square	7.7602
DF	4
Pr > ChiSq	0.1008

Output 36.3.1 shows the frequency table and chi-square test for Region 1. The frequency table lists the variable values (hair color) in the order in which they appear in the data set. The "Test Percent" column lists the hypothesized percentages for the chi-square test. Always check that you have ordered the TESTP= percentages to correctly match the order of the variable levels.

Output 36.3.2 shows the deviation plot for Region 1, which displays the relative deviations from the hypothesized values. The relative deviation for a level is the difference between the observed and hypothesized (test) percentage divided by the test percentage. You can suppress the chi-square *p*-value that is displayed by default in the deviation plot by specifying the NOSTATS *plot-option*.

Output 36.3.2 Deviation Plot for Region 1



Output 36.3.3 and Output 36.3.4 show the results for Region 2. PROC FREQ computes a chi-square statistic for each region. The chi-square statistic is significant at the 0.05 level for Region 2 (p=0.0003) but not for Region 1. This indicates a significant departure from the hypothesized percentages in Region 2.

Output 36.3.3 Frequency Table and Chi-Square Test for Region 2

Hair Color of European Children

Geographic Region=2

The FREQ Procedure

Hair Color

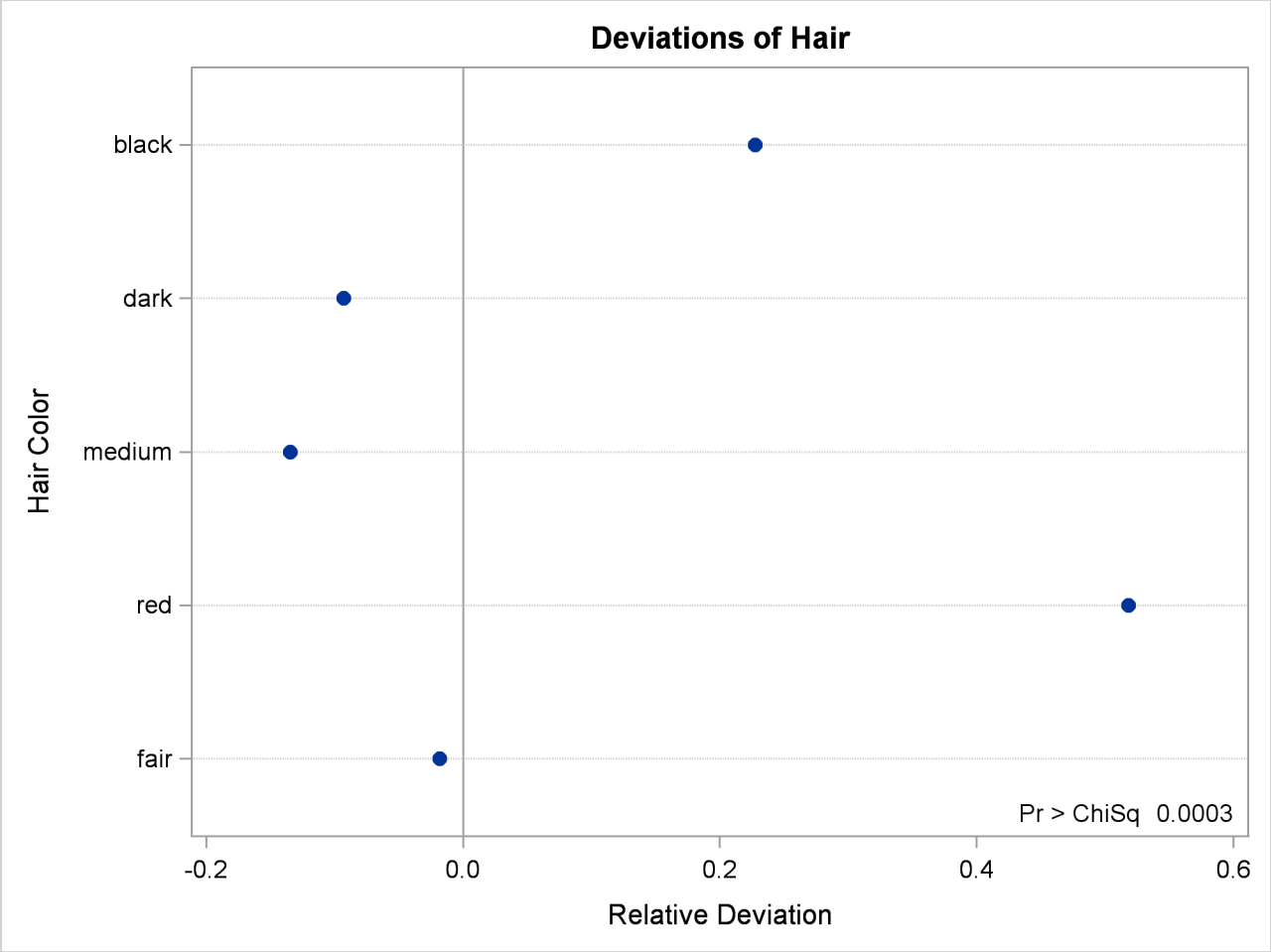
Hair	Frequency	Percent	Test Percent
fair	152	29.46	30.00
red	94	18.22	12.00
medium	134	25.97	30.00
dark	117	22.67	25.00
black	19	3.68	3.00

Geographic Region=2

Chi-Square Test for Specified Proportions

Chi-Square	21.3824
DF	4
Pr > ChiSq	0.0003

Output 36.3.4 Deviation Plot for Region 2



Example 36.4: Binomial Proportions

In this example, PROC FREQ computes binomial proportions, confidence limits, and tests. The example uses the eye and hair color data from Example 36.1. By default, PROC FREQ computes the binomial proportion as the proportion of observations in the first level of the one-way table. You can designate a different level by using the LEVEL= *binomial-option*.

The following PROC FREQ statements compute the proportion of children with brown eyes (from the data set in Example 36.1) and test the null hypothesis that the population proportion equals 50%. These statements also compute an equivalence for the proportion of children with fair hair.

The first TABLES statement requests a one-way frequency table for the variable Eyes. The BINOMIAL option requests the binomial proportion, confidence limits, and test. PROC FREQ computes the proportion with Eyes = 'brown', which is the first level displayed in the table. The AC, WILSON, and EXACT *binomial-options* request the following confidence limits types: Agresti-Coull, Wilson (score), and exact (Clopper-Pearson). By default, PROC FREQ provides Wald and exact (Clopper-Pearson) confidence limits for the binomial proportion. The BINOMIAL option also produces an asymptotic Wald test that the proportion equals 0.5. You can specify a different test proportion with the P= *binomial-option*. The ALPHA=0.1 option specifies that $\alpha = 10\%$, which produces 90% confidence limits.

The second TABLES statement requests a one-way frequency table for the variable Hair. The BINOMIAL option requests the proportion for the first level, Hair = 'fair'. The EQUIV *binomial-option* requests an equivalence test for the binomial proportion. The P=.28 option specifies 0.28 as the null hypothesis proportion, and the MARGIN=.1 option specifies 0.1 as the equivalence test margin.

```
proc freq data=Color order=freq;
   tables Eyes / binomial(ac wilson exact) alpha=.1;
   tables Hair / binomial(equiv p=.28 margin=.1);
   weight Count;
   title 'Hair and Eye Color of European Children';
run;
```

Output 36.4.1 displays the results for eye color, and Output 36.4.2 displays the results for hair color.

Output 36.4.1 Binomial Proportion for Eye Color

Hair and Eye Color of European Children

The FREQ Procedure

Eye Color

Eyes	Frequency	Percent	Cumulative Frequency	Cumulative Percent
brown	341	44.75	341	44.75
blue	222	29.13	563	73.88
green	199	26.12	762	100.00

Binomial Proportion for Eyes = brown

Proportion	0.4475
ASE	0.0180

Type	90% Confidence Limits	
Wilson	0.4181	0.4773
Agresti-Coull	0.4181	0.4773
Clopper-Pearson (Exact)	0.4174	0.4779

Test of H0: Proportion = 0.5

ASE under H0	0.0181
Z	-2.8981
One-sided Pr < Z	0.0019
Two-sided Pr > |Z|	0.0038

The frequency table in Output 36.4.1 displays the values of Eyes in order of descending frequency count. PROC FREQ computes the proportion of children in the first level displayed in the frequency table, Eyes = 'brown'. Output 36.4.1 displays the binomial proportion confidence limits and test. The confidence limits are 90% confidence limits. If you do not specify the ALPHA= option, PROC FREQ computes 95% confidence limits by default. Because the value of Z is less than zero, PROC FREQ displays the a left-sided

p-value (0.0019). This small p-value supports the alternative hypothesis that the true value of the proportion of children with brown eyes is less than 50%.

Output 36.4.2 displays the equivalence test results produced by the second TABLES statement. The null hypothesis proportion is 0.28 and the equivalence margins are –0.1 and 0.1, which yield equivalence limits of 0.18 and 0.38. PROC FREQ provides two one-sided tests (TOST) for equivalence. The small p-value indicates rejection of the null hypothesis in favor of the alternative that the proportion is equivalent to the null value.

Output 36.4.2 Binomial Proportion for Hair Color

Hair Color

Hair	Frequency	Percent	Cumulative Frequency	Cumulative Percent
fair	228	29.92	228	29.92
medium	217	28.48	445	58.40
dark	182	23.88	627	82.28
red	113	14.83	740	97.11
black	22	2.89	762	100.00

Equivalence Analysis

H0: P − p0 <= Lower Margin or >= Upper Margin
Ha: Lower Margin < P − p0 < Upper Margin

p0 = 0.28 Lower Margin = −0.1 Upper Margin = 0.1

Proportion	ASE (Sample)
0.2992	0.0166

Two One-Sided Tests (TOST)

Test	Z	P-Value	
Lower Margin	7.1865	Pr > Z	<.0001
Upper Margin	−4.8701	Pr < Z	<.0001
Overall			<.0001

Equivalence Limits		90% Confidence Limits	
0.1800	0.3800	0.2719	0.3265

Example 36.5: Analysis of a 2x2 Contingency Table

This example computes chi-square tests and Fisher's exact test to compare the probability of coronary heart disease for two types of diet. It also estimates the relative risks and computes exact confidence limits for the odds ratio.

The data set FatComp contains hypothetical data for a case-control study of high fat diet and the risk of coronary heart disease. The data are recorded as cell counts, where the variable Count contains the frequencies for each exposure and response combination. The data set is sorted in descending order by the variables Exposure and Response, so that the first cell of the 2×2 table contains the frequency of positive exposure and positive response. The FORMAT procedure creates formats to identify the type of exposure and response with character values.

```
proc format;
   value ExpFmt 1='High Cholesterol Diet'
                0='Low Cholesterol Diet';
   value RspFmt 1='Yes'
                0='No';
run;

data FatComp;
   input Exposure Response Count;
   label Response='Heart Disease';
   datalines;
0 0  6
0 1  2
1 0  4
1 1 11
;

proc sort data=FatComp;
   by descending Exposure descending Response;
run;
```

In the following PROC FREQ statements, ORDER=DATA option orders the contingency table values by their order in the input data set. The TABLES statement requests a two-way table of Exposure by Response. The CHISQ option produces several chi-square tests, while the RELRISK option produces relative risk measures. The EXACT statement requests the exact Pearson chi-square test and exact confidence limits for the odds ratio.

```
proc freq data=FatComp order=data;
   format Exposure ExpFmt. Response RspFmt.;
   tables Exposure*Response / chisq relrisk;
   exact pchi or;
   weight Count;
   title 'Case-Control Study of High Fat/Cholesterol Diet';
run;
```

The contingency table in Output 36.5.1 displays the variable values so that the first table cell contains the frequency for the first cell in the data set (the frequency of positive exposure and positive response).

Output 36.5.1 Contingency Table

Case-Control Study of High Fat/Cholesterol Diet

The FREQ Procedure

Table of Exposure by Response

Exposure	Response(Heart Disease)		
Frequency Percent Row Pct Col Pct	Yes	No	Total
High Cholesterol Diet	11 47.83 73.33 84.62	4 17.39 26.67 40.00	15 65.22
Low Cholesterol Diet	2 8.70 25.00 15.38	6 26.09 75.00 60.00	8 34.78
Total	13 56.52	10 43.48	23 100.00

Output 36.5.2 displays the chi-square statistics. Because the expected counts in some of the table cells are small, PROC FREQ gives a warning that the asymptotic chi-square tests might not be appropriate. In this case, the exact tests are appropriate. The alternative hypothesis for this analysis states that coronary heart disease is more likely to be associated with a high fat diet, so a one-sided test is desired. Fisher's exact right-sided test analyzes whether the probability of heart disease in the high fat group exceeds the probability of heart disease in the low fat group; because this p-value is small, the alternative hypothesis is supported.

The odds ratio, displayed in Output 36.5.3, provides an estimate of the relative risk when an event is rare. This estimate indicates that the odds of heart disease is 8.25 times higher in the high fat diet group; however, the wide confidence limits indicate that this estimate has low precision.

Output 36.5.2 Chi-Square Statistics

Statistic	DF	Value	Prob
Chi-Square	1	4.9597	0.0259
Likelihood Ratio Chi-Square	1	5.0975	0.0240
Continuity Adj. Chi-Square	1	3.1879	0.0742
Mantel-Haenszel Chi-Square	1	4.7441	0.0294
Phi Coefficient		0.4644	
Contingency Coefficient		0.4212	
Cramer's V		0.4644	

WARNING: 50% of the cells have expected counts less than 5.
(Asymptotic) Chi-Square may not be a valid test.

Pearson Chi-Square Test

Chi-Square	4.9597
DF	1
Asymptotic Pr > ChiSq	0.0259
Exact Pr >= ChiSq	0.0393

Fisher's Exact Test

Cell (1,1) Frequency (F)	11
Left-sided Pr <= F	0.9967
Right-sided Pr >= F	0.0367
Table Probability (P)	0.0334
Two-sided Pr <= P	0.0393

Output 36.5.3 Relative Risk

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	8.2500	1.1535	59.0029
Cohort (Col1 Risk)	2.9333	0.8502	10.1204
Cohort (Col2 Risk)	0.3556	0.1403	0.9009

Odds Ratio (Case-Control Study)

Odds Ratio	8.2500
Asymptotic Conf Limits	
95% Lower Conf Limit	1.1535
95% Upper Conf Limit	59.0029
Exact Conf Limits	
95% Lower Conf Limit	0.8677
95% Upper Conf Limit	105.5488

Example 36.6: Output Data Set of Chi-Square Statistics

This example uses the Color data from Example 36.1 to output the Pearson chi-square and the likelihood-ratio chi-square statistics to a SAS data set. The following PROC FREQ statements create a two-way table of eye color versus hair color.

```
proc freq data=Color order=data;
   tables Eyes*Hair / expected cellchi2 norow nocol chisq;
   output out=ChiSqData n nmiss pchi lrchi;
   weight Count;
   title 'Chi-Square Tests for 3 by 5 Table of Eye and Hair Color';
run;

proc print data=ChiSqData noobs;
   title1 'Chi-Square Statistics for Eye and Hair Color';
   title2 'Output Data Set from the FREQ Procedure';
run;
```

The EXPECTED option displays expected cell frequencies in the crosstabulation table, and the CELLCHI2 option displays the cell contribution to the overall chi-square. The NOROW and NOCOL options suppress the display of row and column percents in the crosstabulation table. The CHISQ option produces chi-square tests.

The OUTPUT statement creates the ChiSqData output data set and specifies the statistics to include. The N option requests the number of nonmissing observations, the NMISS option stores the number of missing observations, and the PCHI and LRCHI options request Pearson and likelihood-ratio chi-square statistics, respectively, together with their degrees of freedom and p-values.

The preceding statements produce Output 36.6.1 and Output 36.6.2. The contingency table in Output 36.6.1 displays eye and hair color in the order in which they appear in the Color data set. The Pearson chi-square statistic in Output 36.6.2 provides evidence of an association between eye and hair color (p=0.0073). The cell chi-square values show that most of the association is due to more green-eyed children with fair or red hair and fewer with dark or black hair. The opposite occurs with the brown-eyed children.

Output 36.6.3 displays the output data set created by the OUTPUT statement. It includes one observation that contains the sample size, the number of missing values, and the chi-square statistics and corresponding degrees of freedom and p-values as in Output 36.6.2.

Output 36.6.1 Contingency Table

Chi-Square Tests for 3 by 5 Table of Eye and Hair Color

The FREQ Procedure

Table of Eyes by Hair

Eyes(Eye Color) Hair(Hair Color)

Frequency / Expected / Cell Chi-Square / Percent	fair	red	medium	dark	black	Total
blue	69	28	68	51	6	222
	66.425	32.921	63.22	53.024	6.4094	
	0.0998	0.7357	0.3613	0.0772	0.0262	
	9.06	3.67	8.92	6.69	0.79	29.13
green	69	38	55	37	0	199
	59.543	29.51	56.671	47.53	5.7454	
	1.5019	2.4422	0.0492	2.3329	5.7454	
	9.06	4.99	7.22	4.86	0.00	26.12
brown	90	47	94	94	16	341
	102.03	50.568	97.109	81.446	9.8451	
	1.4187	0.2518	0.0995	1.935	3.8478	
	11.81	6.17	12.34	12.34	2.10	44.75
Total	228	113	217	182	22	762
	29.92	14.83	28.48	23.88	2.89	100.00

Output 36.6.2 Chi-Square Statistics

Statistic	DF	Value	Prob
Chi-Square	8	20.9248	0.0073
Likelihood Ratio Chi-Square	8	25.9733	0.0011
Mantel-Haenszel Chi-Square	1	3.7838	0.0518
Phi Coefficient		0.1657	
Contingency Coefficient		0.1635	
Cramer's V		0.1172	

Output 36.6.3 Output Data Set

Chi-Square Statistics for Eye and Hair Color
Output Data Set from the FREQ Procedure

N	NMISS	_PCHI_	DF_PCHI	P_PCHI	_LRCHI_	DF_LRCHI	P_LRCHI
762	0	20.9248	8	.007349898	25.9733	8	.001061424

Example 36.7: Cochran-Mantel-Haenszel Statistics

The data set Migraine contains hypothetical data for a clinical trial of migraine treatment. Subjects of both genders receive either a new drug therapy or a placebo. Their response to treatment is coded as 'Better' or 'Same'. The data are recorded as cell counts, and the number of subjects for each treatment and response combination is recorded in the variable Count.

```
data Migraine;
   input Gender $ Treatment $ Response $ Count @@;
   datalines;
female Active  Better 16   female Active  Same 11
female Placebo Better  5   female Placebo Same 20
male   Active  Better 12   male   Active  Same 16
male   Placebo Better  7   male   Placebo Same 19
;
```

The following PROC FREQ statements create a multiway table stratified by Gender, where Treatment forms the rows and Response forms the columns. The RELRISK option in the TABLES statement requests the odds ratio and relative risks for the two-way tables of Treatment by Response. The PLOTS= option requests a relative risk plot, which shows the relative risk and its confidence limits for each level of Gender. The CMH option requests Cochran-Mantel-Haenszel statistics for the multiway table. For this stratified 2×2 table, the CMH option also produces estimates of the common relative risk and the Breslow-Day test for homogeneity of the odds ratios. The NOPRINT option suppresses the display of the crosstabulation tables.

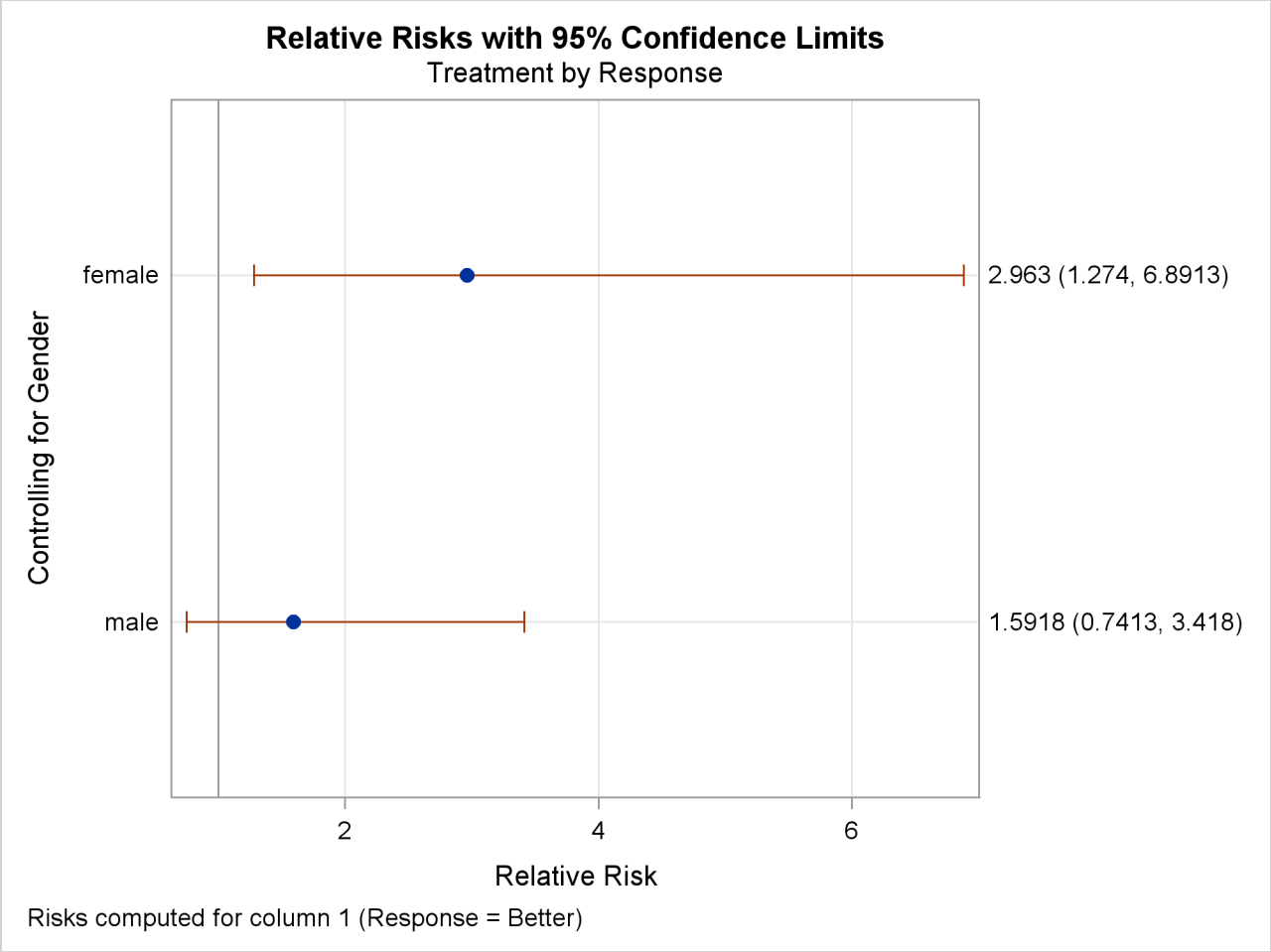
```
ods graphics on;
proc freq data=Migraine;
   tables Gender*Treatment*Response /
          relrisk plots(only)=relriskplot(stats) cmh noprint;
   weight Count;
   title 'Clinical Trial for Treatment of Migraine Headaches';
run;
ods graphics off;
```

Output 36.7.1 through Output 36.7.4 show the results of the analysis. The relative risk plot (Output 36.7.1) displays the relative risks and confidence limits for the two levels of Gender. Output 36.7.2 displays the CMH statistics. For a stratified 2×2 table, the three CMH statistics test the same hypothesis. The significant p-value (0.004) indicates that the association between treatment and response remains strong after adjusting for gender.

The CMH option also produces a table of overall relative risks, as shown in Output 36.7.3. Because this is a prospective study, the relative risk estimate assesses the effectiveness of the new drug; the "Cohort (Col1 Risk)" values are the appropriate estimates for the first column (the risk of improvement). The probability of migraine improvement with the new drug is just over two times the probability of improvement with the placebo.

The large p-value for the Breslow-Day test (0.2218) in Output 36.7.4 indicates no significant gender difference in the odds ratios.

Output 36.7.1 Relative Risk Plot



Output 36.7.2 Cochran-Mantel-Haenszel Statistics

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	8.3052	0.0040
2	Row Mean Scores Differ	1	8.3052	0.0040
3	General Association	1	8.3052	0.0040

Output 36.7.3 CMH Option: Common Relative Risks

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	Mantel-Haenszel	3.3132	1.4456	7.5934
	Logit	3.2941	1.4182	7.6515
Cohort (Col1 Risk)	Mantel-Haenszel	2.1636	1.2336	3.7948
	Logit	2.1059	1.1951	3.7108
Cohort (Col2 Risk)	Mantel-Haenszel	0.6420	0.4705	0.8761
	Logit	0.6613	0.4852	0.9013

Output 36.7.4 CMH Option: Breslow-Day Test

Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square	1.4929
DF	1
Pr > ChiSq	0.2218

Example 36.8: Cochran-Armitage Trend Test

The data set Pain contains hypothetical data for a clinical trial of a drug therapy to control pain. The clinical trial investigates whether adverse responses increase with larger drug doses. Subjects receive either a placebo or one of four drug doses. An adverse response is recorded as Adverse='Yes'; otherwise, it is recorded as Adverse='No'. The number of subjects for each drug dose and response combination is contained in the variable Count.

```
data pain;
   input Dose Adverse $ Count @@;
   datalines;
0 No 26   0 Yes  6
1 No 26   1 Yes  7
2 No 23   2 Yes  9
3 No 18   3 Yes 14
4 No  9   4 Yes 23
;
```

The following PROC FREQ statements provide a trend analysis. The TABLES statement requests a table of Adverse by Dose. The MEASURES option produces measures of association, and the CL option produces confidence limits for these measures. The TREND option tests for a trend across the ordinal values of the variable Dose with the Cochran-Armitage test. The EXACT statement produces exact p-values for this test, and the MAXTIME= option terminates the exact computations if they do not complete within 60 seconds. The TEST statement computes an asymptotic test for Somers' $D(R|C)$.

The PLOTS= option requests a frequency plot for the table of Adverse by Dose. By default, PROC FREQ provides a bar chart for the frequency plot. The TWOWAY=STACKED option requests a stacked layout, where the bars correspond to the column variable (Dose) values, and the row variable (Adverse) frequencies are stacked within each bar.

```
ods graphics on;
proc freq data=Pain;
   tables Adverse*Dose / trend measures cl
          plots=freqplot(twoway=stacked orient=horizontal scale=percent);
   test smdrc;
   exact trend / maxtime=60;
   weight Count;
   title 'Clinical Trial for Treatment of Pain';
run;
ods graphics off;
```

Output 36.8.1 through Output 36.8.4 display the results of the analysis. The "Col Pct" values in Output 36.8.1 show the expected increasing trend in the proportion of adverse effects with the increasing dosage (from 18.75% to 71.88%). The corresponding bar chart (Output 36.8.2) also shows this increasing trend.

Output 36.8.1 Contingency Table

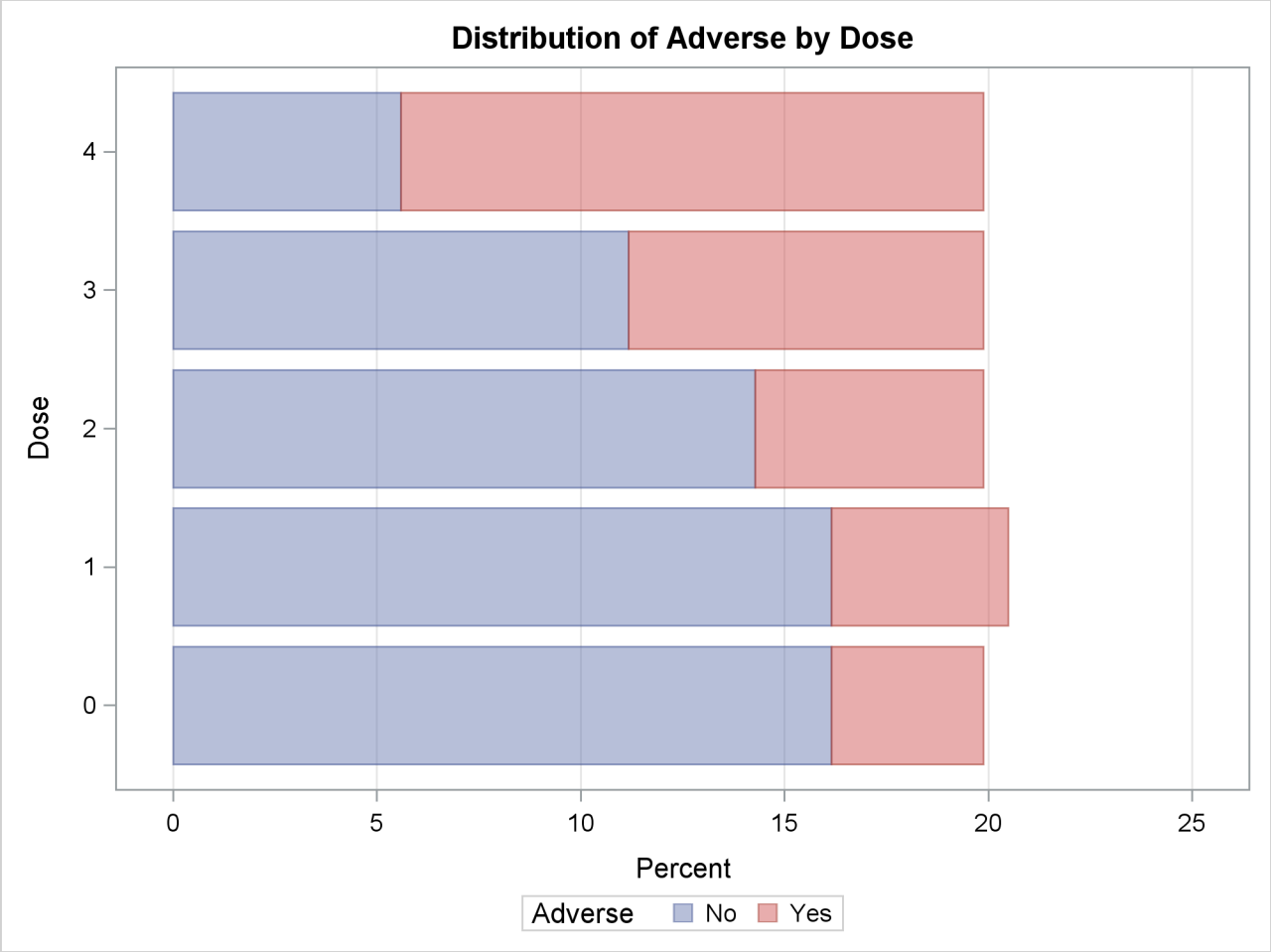
Clinical Trial for Treatment of Pain

The FREQ Procedure

Table of Adverse by Dose

Adverse	Dose					
Frequency Percent Row Pct Col Pct	0	1	2	3	4	Total
No	26 16.15 25.49 81.25	26 16.15 25.49 78.79	23 14.29 22.55 71.88	18 11.18 17.65 56.25	9 5.59 8.82 28.13	102 63.35
Yes	6 3.73 10.17 18.75	7 4.35 11.86 21.21	9 5.59 15.25 28.13	14 8.70 23.73 43.75	23 14.29 38.98 71.88	59 36.65
Total	32 19.88	33 20.50	32 19.88	32 19.88	32 19.88	161 100.00

Output 36.8.2 Stacked Bar Chart of Percents



Output 36.8.3 displays the measures of association produced by the MEASURES option. Somers' $D(R|C)$ measures the association treating the row variable (Adverse) as the response and the column variable (Dose) as a predictor. Because the asymptotic 95% confidence limits do not contain zero, this indicates a strong positive association. Similarly, the Pearson and Spearman correlation coefficients show evidence of a strong positive association, as hypothesized.

The Cochran-Armitage test (Output 36.8.4) supports the trend hypothesis. The small left-sided p-values for the Cochran-Armitage test indicate that the probability of the Row 1 level (Adverse='No') decreases as Dose increases or, equivalently, that the probability of the Row 2 level (Adverse='Yes') increases as Dose increases. The two-sided p-value tests against either an increasing or decreasing alternative. This is an appropriate hypothesis when you want to determine whether the drug has progressive effects on the probability of adverse effects but the direction is unknown.

Output 36.8.3 Measures of Association

Statistic	Value	ASE	95% Confidence Limits	
Gamma	0.5313	0.0935	0.3480	0.7146
Kendall's Tau-b	0.3373	0.0642	0.2114	0.4631
Stuart's Tau-c	0.4111	0.0798	0.2547	0.5675
Somers' D C|R	0.4427	0.0837	0.2786	0.6068
Somers' D R|C	0.2569	0.0499	0.1592	0.3547
Pearson Correlation	0.3776	0.0714	0.2378	0.5175
Spearman Correlation	0.3771	0.0718	0.2363	0.5178
Lambda Asymmetric C|R	0.1250	0.0662	0.0000	0.2547
Lambda Asymmetric R|C	0.2373	0.0837	0.0732	0.4014
Lambda Symmetric	0.1604	0.0621	0.0388	0.2821
Uncertainty Coefficient C|R	0.0515	0.0191	0.0140	0.0890
Uncertainty Coefficient R|C	0.1261	0.0467	0.0346	0.2175
Uncertainty Coefficient Symmetric	0.0731	0.0271	0.0199	0.1262

Somers' D R|C	
Somers' D R|C	0.2569
ASE	0.0499
95% Lower Conf Limit	0.1592
95% Upper Conf Limit	0.3547

Test of H0: Somers' D R|C = 0

ASE under H0	0.0499
Z	5.1511
One-sided Pr > Z	<.0001
Two-sided Pr > |Z|	<.0001

Output 36.8.4 Trend Test

Cochran-Armitage Trend Test	
Statistic (Z)	-4.7918
Asymptotic Test	
One-sided Pr < Z	<.0001
Two-sided Pr > |Z|	<.0001
Exact Test	
One-sided Pr <= Z	7.237E-07
Two-sided Pr >= |Z|	1.324E-06

Example 36.9: Friedman's Chi-Square Test

Friedman's test is a nonparametric test for treatment differences in a randomized complete block design. Each block of the design might be a subject or a homogeneous group of subjects. If blocks are groups of subjects, the number of subjects in each block must equal the number of treatments. Treatments are randomly assigned to subjects within each block. If there is one subject per block, then the subjects are repeatedly measured once under each treatment. The order of treatments is randomized for each subject.

In this setting, Friedman's test is identical to the ANOVA (row means scores) CMH statistic when the analysis uses rank scores (SCORES=RANK). The three-way table uses subject (or subject group) as the stratifying variable, treatment as the row variable, and response as the column variable. PROC FREQ handles ties by assigning midranks to tied response values. If there are multiple subjects per treatment in each block, the ANOVA CMH statistic is a generalization of Friedman's test.

The data set Hypnosis contains data from a study investigating whether hypnosis has the same effect on skin potential (measured in millivolts) for four emotions (Lehmann 1975, p. 264). Eight subjects are asked to display fear, joy, sadness, and calmness under hypnosis. The data are recorded as one observation per subject for each emotion.

```
data Hypnosis;
   length Emotion $ 10;
   input Subject Emotion $ SkinResponse @@;
   datalines;
1 fear 23.1  1 joy 22.7  1 sadness 22.5  1 calmness 22.6
2 fear 57.6  2 joy 53.2  2 sadness 53.7  2 calmness 53.1
3 fear 10.5  3 joy  9.7  3 sadness 10.8  3 calmness  8.3
4 fear 23.6  4 joy 19.6  4 sadness 21.1  4 calmness 21.6
5 fear 11.9  5 joy 13.8  5 sadness 13.7  5 calmness 13.3
6 fear 54.6  6 joy 47.1  6 sadness 39.2  6 calmness 37.0
7 fear 21.0  7 joy 13.6  7 sadness 13.7  7 calmness 14.8
8 fear 20.3  8 joy 23.6  8 sadness 16.3  8 calmness 14.8
;
```

In the following PROC FREQ statements, the TABLES statement creates a three-way table stratified by Subject and a two-way table; the variables Emotion and SkinResponse form the rows and columns of each table. The CMH2 option produces the first two Cochran-Mantel-Haenszel statistics, the option SCORES=RANK specifies that rank scores are used to compute these statistics, and the NOPRINT option suppresses the contingency tables. These statements produce Output 36.9.1 and Output 36.9.2.

```
proc freq data=Hypnosis;
   tables Subject*Emotion*SkinResponse /
          cmh2 scores=rank noprint;
run;

proc freq data=Hypnosis;
   tables Emotion*SkinResponse /
          cmh2 scores=rank noprint;
run;
```

Because the CMH statistics in Output 36.9.1 are based on rank scores, the Row Mean Scores Differ statistic is identical to Friedman's chi-square ($Q = 6.45$). The p-value of 0.0917 indicates that differences in skin potential response for different emotions are significant at the 10% level but not at the 5% level.

When you do not stratify by subject, the Row Mean Scores Differ CMH statistic is identical to a Kruskal-Wallis test and is not significant (p=0.9038 in Output 36.9.2). Thus, adjusting for subject is critical to reducing the background variation due to subject differences.

Output 36.9.1 CMH Statistics: Stratifying by Subject

Clinical Trial for Treatment of Pain

The FREQ Procedure

Summary Statistics for Emotion by SkinResponse
Controlling for Subject

Cochran-Mantel-Haenszel Statistics (Based on Rank Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	0.2400	0.6242
2	Row Mean Scores Differ	3	6.4500	0.0917

Output 36.9.2 CMH Statistics: No Stratification

Clinical Trial for Treatment of Pain

The FREQ Procedure

Summary Statistics for Emotion by SkinResponse

Cochran-Mantel-Haenszel Statistics (Based on Rank Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	0.0001	0.9933
2	Row Mean Scores Differ	3	0.5678	0.9038

Example 36.10: Cochran's Q Test

When a binary response is measured several times or under different conditions, Cochran's Q tests that the marginal probability of a positive response is unchanged across the times or conditions. When there are more than two response categories, you can use the CATMOD procedure to fit a repeated-measures model.

The data set Drugs contains data for a study of three drugs to treat a chronic disease (Agresti 2002). Forty-six subjects receive drugs A, B, and C. The response to each drug is either favorable ('F') or unfavorable ('U').

```
proc format;
   value $ResponseFmt 'F'='Favorable'
                      'U'='Unfavorable';
run;

data drugs;
   input Drug_A $ Drug_B $ Drug_C $ Count @@;
   datalines;
F F F  6   U F F  2
F F U 16   U F U  4
F U F  2   U U F  6
F U U  4   U U U  6
;
```

The following statements create one-way frequency tables of the responses to each drug. The AGREE option produces Cochran's Q and other measures of agreement for the three-way table. These statements produce Output 36.10.1 through Output 36.10.5.

```
proc freq data=Drugs;
   tables Drug_A Drug_B Drug_C / nocum;
   tables Drug_A*Drug_B*Drug_C / agree noprint;
   format Drug_A Drug_B Drug_C $ResponseFmt.;
   weight Count;
   title 'Study of Three Drug Treatments for a Chronic Disease';
run;
```

The one-way frequency tables in Output 36.10.1 provide the marginal response for each drug. For drugs A and B, 61% of the subjects reported a favorable response while 35% of the subjects reported a favorable response to drug C. Output 36.10.2 and Output 36.10.3 display measures of agreement for the 'Favorable' and 'Unfavorable' levels of drug A, respectively. McNemar's test shows a strong discordance between drugs B and C when the response to drug A is favorable.

Output 36.10.1 One-Way Frequency Tables

Study of Three Drug Treatments for a Chronic Disease

The FREQ Procedure

Drug_A	Frequency	Percent
Favorable	28	60.87
Unfavorable	18	39.13

Drug_B	Frequency	Percent
Favorable	28	60.87
Unfavorable	18	39.13

Drug_C	Frequency	Percent
Favorable	16	34.78
Unfavorable	30	65.22

Output 36.10.2 Measures of Agreement for Drug A Favorable

McNemar's Test	
Statistic (S)	10.8889
DF	1
Pr > S	0.0010

Simple Kappa Coefficient	
Kappa	−0.0328
ASE	0.1167
95% Lower Conf Limit	−0.2615
95% Upper Conf Limit	0.1960

Output 36.10.3 Measures of Agreement for Drug A Unfavorable

McNemar's Test	
Statistic (S)	0.4000
DF	1
Pr > S	0.5271

Simple Kappa Coefficient	
Kappa	−0.1538
ASE	0.2230
95% Lower Conf Limit	−0.5909
95% Upper Conf Limit	0.2832

Output 36.10.4 displays the overall kappa coefficient. The small negative value of kappa indicates no agreement between drug B response and drug C response.

Cochran's Q is statistically significant (p=0.0144 in Output 36.10.5), which leads to rejection of the hypothesis that the probability of favorable response is the same for the three drugs.

Output 36.10.4 Overall Measures of Agreement

Overall Kappa Coefficient	
Kappa	−0.0588
ASE	0.1034
95% Lower Conf Limit	−0.2615
95% Upper Conf Limit	0.1439

Test for Equal Kappa Coefficients	
Chi-Square	0.2314
DF	1
Pr > ChiSq	0.6305

Output 36.10.5 Cochran's *Q* Test

Cochran's Q, for Drug_A by Drug_B by Drug_C	
Statistic (Q)	8.4706
DF	2
Pr > Q	0.0145

References

Agresti, A. (1992), "A Survey of Exact Inference for Contingency Tables," *Statistical Science*, 7(1), 131–177.

Agresti, A. (2002), *Categorical Data Analysis*, Second Edition, New York: John Wiley & Sons.

Agresti, A. (2007), *An Introduction to Categorical Data Analysis*, Second Edition, New York: John Wiley & Sons.

Agresti, A. and Coull, B. A. (1998), "Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions," *The American Statistician*, 52, 119–126.

Agresti, A., Mehta, C. R., and Patel, N. R. (1990), "Exact Inference for Contingency Tables with Ordered Categories," *Journal of the American Statistical Association*, 85, 453–458.

Agresti, A. and Min, Y. (2001), "On Small-Sample Confidence Intervals for Parameters in Discrete Distributions," *Biometrics*, 57, 963–971.

Agresti, A., Wackerly, D., and Boyett, J. M. (1979), "Exact Conditional Tests for Cross-Classifications: Approximation of Attained Significance Levels," *Psychometrika*, 44, 75–83.

Bangdiwala, S. I. (1988), "The Agreement Chart," Institute of Statistics Mimeo Series No. 1859, Department of Biostatistics, University of North Carolina at Chapel Hill.

Bangdiwala, S. I. and Bryan, H. E. (1987), "Using SAS Software Graphical Procedures for the Observer Agreement Chart," in *Proceedings of the Twelfth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc., 1083–1088.

Bangdiwala, S. I., Haedo, A. S., Natal, M. L., and Villaveces, A. (2008), "The Agreement Chart as an Alternative to the Receiver-Operating Characteristic Curve for Diagnostic Tests," *Journal of Clinical Epidemiology*, 61, 866–874.

Barker, L., Rolka, H., Rolka, D., and Brown, C. (2001), "Equivalence Testing for Binomial Random Variables: Which Test to Use?," *The American Statistician*, 55, 279–287.

Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, Second Edition, New York: Springer-Verlag.

Birch, M. W. (1965), "The Detection of Partial Association, II: The General Case," *Journal of the Royal Statistical Society, B*, 27, 111–124.

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

Bowker, A. H. (1948), "Bowker's Test for Symmetry," *Journal of the American Statistical Association*, 43, 572–574.

Breslow, N. E. (1996), "Statistics in Epidemiology: The Case-Control Study," *Journal of the American Statistical Association*, 91, 14–26.

Breslow, N. E. and Day, N. E. (1980), *Statistical Methods in Cancer Research, Volume I: The Analysis of Case-Control Studies*, IARC Scientific Publications, No. 32, Lyon, France: International Agency for Research on Cancer.

Breslow, N. E. and Day, N. E. (1987), *Statistical Methods in Cancer Research, Volume II: The Design and Analysis of Cohort Studies*, IARC Scientific Publications, No. 82, Lyon, France: International Agency for Research on Cancer.

Bross, I. D. J. (1958), "How to Use Ridit Analysis," *Biometrics*, 14, 18–38.

Brown, L. D., Cai, T. T., and DasGupta, A. (2001), "Interval Estimation for a Binomial Proportion," *Statistical Science* 16, 101–133.

Brown, M. B. and Benedetti, J. K. (1977), "Sampling Behavior of Tests for Correlation in Two-Way Contingency Tables," *Journal of the American Statistical Association*, 72, 309–315.

Chan, I. S. F. (1998), "Exact Tests of Equivalence and Efficacy with a Non-Zero Lower Bound for Comparative Studies," *Statistics in Medicine*, 17, 1403–1413.

Chan, I. S. F. (2003), "Proving Non-Inferiority or Equivalence of Two Treatments with Dichotomous Endpoints Using Exact Methods," *Statistical Methods in Medical Research*, 12, 37–58.

Chan, I. S. F. and Zhang, Z. (1999), "Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions," *Biometrics*, 55, 1202–1209.

Chow, S., Shao, J., and Wang, H. (2003), *Sample Size Calculations in Clinical Research*, Boca Raton, FL: CRC Press.

Cicchetti, D. V. and Allison, T. (1971), "A New Procedure for Assessing Reliability of Scoring EEG Sleep Recordings," *American Journal of EEG Technology*, 11, 101–109.

Clopper, C. J., and Pearson, E. S. (1934), "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial," *Biometrika* 26, 404–413.

Cochran, W. G. (1950), "The Comparison of Percentages in Matched Samples," *Biometrika*, 37, 256–266.

Cochran, W. G. (1954), "Some Methods for Strengthening the Common χ^2 Tests," *Biometrics*, 10, 417–451.

Collett, D. (1991), *Modelling Binary Data*, London: Chapman & Hall.

Cohen, J. (1960), "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, 20, 37–46.

Dimitrienko, A., Molenberghs, G., Chuang-Stein, C., and Offen, W. (2005), *Analysis of Clinical Trials Using SAS: A Practical Guide*, Cary, NC: SAS Institute Inc.

Drasgow, F. (1986), "Polychoric and Polyserial Correlations" in *Encyclopedia of Statistical Sciences*, vol. 7, ed. S. Kotz and N. L. Johnson, New York: John Wiley & Sons, 68–74.

Dunnett, C. W., and Gent, M. (1977), "Significance Testing to Establish Equivalence Between Treatments, with Special Reference to Data in the Form of 2×2 Tables," *Biometrics*, 33, 593–602.

Farrington, C. P., and Manning, G. (1990), "Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-Zero Risk Difference or Non-Unity Relative Risk," *Statistics in Medicine*, 9, 1447–1454.

Fienberg, S. E. (1980), *The Analysis of Cross-Classified Data*, Second Edition, Cambridge, MA: MIT Press.

Fleiss, J. L., Levin, B., and Paik, M. C. (2003), *Statistical Methods for Rates and Proportions*, Third Edition, New York: John Wiley & Sons.

Fleiss, J. L. and Cohen, J. (1973), "The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability," *Educational and Psychological Measurement*, 33, 613–619.

Fleiss, J. L., Cohen, J., and Everitt, B. S. (1969), "Large-Sample Standard Errors of Kappa and Weighted Kappa," *Psychological Bulletin*, 72, 323–327.

Freeman, G. H. and Halton, J. H. (1951), "Note on an Exact Treatment of Contingency, Goodness of Fit and Other Problems of Significance," *Biometrika*, 38, 141–149.

Friendly, M. (2000), *Visualizing Categorical Data*, Cary, NC: SAS Institute Inc.

Gail, M. and Mantel, N. (1977), "Counting the Number of $r \times c$ Contingency Tables with Fixed Margins," *Journal of the American Statistical Association*, 72, 859–862.

Gail, M. and Simon, R. (1985), "Tests for Qualitative Interactions between Treatment Effects and Patient Subsets," *Biometrics*, 41, 361–372.

Gart, J. J. (1971), "The Comparison of Proportions: A Review of Significance Tests, Confidence Intervals, and Adjustments for Stratification," *Review of the International Statistical Institute*, 39(2), 148–169.

Gart, J. J. and Nam, J. (1988), "Approximate Interval Estimation of the Ratio of Binomial Parameters: A Review and Corrections for Skewness," *Biometrics*, 44, 323–338.

Goodman, L. A. and Kruskal, W. H. (1979), *Measures of Association for Cross Classification*, New York: Springer-Verlag.

Greenland, S. and Robins, J. M. (1985), "Estimators of the Mantel-Haenszel Variance Consistent in Both Sparse Data and Large-Strata Limiting Models," *Biometrics*, 42, 311–323.

Haldane, J. B. S. (1955), "The Estimation and Significance of the Logarithm of a Ratio of Frequencies," *Annals of Human Genetics*, 20, 309–314.

Hauck, W. W. and Anderson, S. (1986), "A Comparison of Large-Sample Confidence Interval Methods for the Difference of Two Binomial Probabilities," *The American Statistician*, 40, 318–322.

Hirji, K. F. (2006), *Exact Analysis of Discrete Data*, Boca Raton, FL: Chapman & Hall/CRC.

Hirji, K. F., Vollset, S. E., Reis, I. M., and Afifi, A. A. (1996), "Exact Tests for Interaction in Several 2×2 Tables," *Journal of Computational and Graphical Statistics*, 5, 209–224.

Hollander, M. and Wolfe, D. A. (1999), *Nonparametric Statistical Methods*, Second Edition, New York: John Wiley & Sons.

Jones, M. P., O'Gorman, T. W., Lemka, J. H., and Woolson, R. F. (1989), "A Monte Carlo Investigation of Homogeneity Tests of the Odds Ratio Under Various Sample Size Configurations," *Biometrics*, 45, 171–181.

Kendall, M. (1955), *Rank Correlation Methods*, Second Edition, London: Charles Griffin and Co.

Kendall, M. and Stuart, A. (1979), *The Advanced Theory of Statistics*, vol. 2, New York: Macmillan.

Kleinbaum, D. G., Kupper, L. L., and Morgenstern, H. (1982), *Epidemiologic Research: Principles and Quantitative Methods*, Research Methods Series, New York: Van Nostrand Reinhold.

Landis, R. J., Heyman, E. R., and Koch, G. G. (1978), "Average Partial Association in Three-way Contingency Tables: A Review and Discussion of Alternative Tests," *International Statistical Review*, 46, 237–254.

Leemis, L. M. and Trivedi, K. S. (1996), "A Comparison of Approximate Interval Estimators for the Bernoulli Parameter," *The American Statistician*, 50, 63–68.

Lehmann, E. L. (1975), *Nonparametrics: Statistical Methods Based on Ranks*, San Francisco: Holden-Day.

Liebetrau, A. M. (1983), *Measures of Association, Quantitative Application in the Social Sciences*, vol. 32, Beverly Hills: Sage Publications.

Mack, G. A. and Skillings, J. H. (1980), "A Friedman-Type Rank Test for Main Effects in a Two-Factor ANOVA," *Journal of the American Statistical Association*, 75, 947–951.

Mantel, N. (1963), "Chi-square Tests with One Degree of Freedom: Extensions of the Mantel-Haenszel Procedure," *Journal of the American Statistical Association*, 58, 690–700.

Mantel, N. and Fleiss, J. L. (1980), "Minimum Expected Cell Size Requirements for the Mantel-Haenszel One-Degree-of-Freedom Chi-Square Test and a Related Rapid Procedure," *American Journal of Epidemiology*, 112, 129–134.

Mantel, N. and Haenszel, W. (1959), "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease," *Journal of the National Cancer Institute*, 22, 719–748.

Margolin, B. H. (1988), "Test for Trend in Proportions," in *Encyclopedia of Statistical Sciences*, vol. 9, ed. S. Kotz and N. L. Johnson, New York: John Wiley & Sons, 334–336.

McNemar, Q. (1947), "Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages," *Psychometrika*, 12, 153–157.

Mehta, C. R. and Patel, N. R. (1983), "A Network Algorithm for Performing Fisher's Exact Test in $r \times c$ Contingency Tables," *Journal of the American Statistical Association*, 78, 427–434.

Mehta, C. R., Patel, N. R., and Gray, R. (1985), "On Computing an Exact Confidence Interval for the Common Odds Ratio in Several 2×2 Contingency Tables," *Journal of the American Statistical Association*, 80, 969–973.

Mehta, C. R., Patel, N. R., and Senchaudhuri, P. (1991), "Exact Stratified Linear Rank Tests for Binary Data," *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, ed. E.M. Keramidas, 200–207.

Mehta, C. R., Patel, N. R., and Tsiatis, A. A. (1984), "Exact Significance Testing to Establish Treatment Equivalence with Ordered Categorical Data," *Biometrics*, 40, 819–825.

Mietinnen, O. and Nurminen, M. (1985), "Comparative Analysis of Two Rates," *Statistics in Medicine*, 4, 213–226.

Narayanan, A. and Watts, D. (1996), "Exact Methods in the NPAR1WAY Procedure," in *Proceedings of the Twenty-First Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc., 1290–1294.

Newcombe, R. G. (1998), "Two-sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods," *Statistics in Medicine*, 17, 857–872.

Newcombe, R. G. (1998), "Interval Estimation for the Difference Between Independent Proportions: Comparison of Eleven Methods," *Statistics in Medicine*, 17, 873–890.

Olsson, U. (1979), "Maximum Likelihood Estimation of the Polychoric Correlation Coefficient," *Psychometrika*, 12, 443–460.

Pirie, W. (1983), "Jonckheere Tests for Ordered Alternatives," in *Encyclopedia of Statistical Sciences*, vol. 4, ed. S. Kotz and N. L. Johnson, New York: John Wiley & Sons, 315–318.

Radlow, R. and Alf, E. F. (1975), "An Alternate Multinomial Assessment of the Accuracy of the Chi-Square Test of Goodness of Fit," *Journal of the American Statistical Association*, 70, 811–813.

Robins, J. M., Breslow, N., and Greenland, S. (1986), "Estimators of the Mantel-Haenszel Variance Consistent in Both Sparse Data and Large-Strata Limiting Models," *Biometrics*, 42, 311–323.

Santner, T. J., Pradhan, V., Senchaudhuri, P., Mehta, C. R., and Tamhane, A. (2007), "Small-Sample Comparisons of Confidence Intervals for the Difference of Two Independent Binomial Proportions," *Computational Statistics and Data Analysis*, 51, 5791–5799.

Santner, T. J. and Snell, M. K. (1980), "Small-Sample Confidence Intervals for $p_1 - p_2$ and p_1/p_2 in 2×2 Contingency Tables," *Journal of the American Statistical Association*, 75, 386–394.

Schuirmann, D. J. (1987), "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability," *Journal of Pharmacokinetics and Biopharmaceutics*, 15, 657–680.

Schuirmann, D. J. (1999), "Confidence Interval Methods for Bioequivalence Testing with Binomial Endpoints," *Proceedings of the Biopharmaceutical Section, ASA*, 227–232.

Silvapulle, M. J. (2001), "Tests Against Qualitative Interaction: Exact Critical Values and Robust Tests," *Biometrics*, 57, 1157–1165.

Snedecor, G. W. and Cochran, W. G. (1989), *Statistical Methods*, Eighth Edition, Ames: Iowa State University Press.

Somers, R. H. (1962), "A New Asymmetric Measure of Association for Ordinal Variables," *American Sociological Review*, 27, 799–811.

Stokes, M. E., Davis, C. S., and Koch, G. G. (2000), *Categorical Data Analysis Using the SAS System*, Second Edition, Cary, NC: SAS Institute Inc.

Tarone, R. E. (1985), "On Heterogeneity Tests Based on Efficient Scores," *Biometrika*, 72, 1, 91–95.

Theil, H. (1972), *Statistical Decomposition Analysis*, Amsterdam: North-Holland Publishing Company.

Thomas, D. G. (1971), "Algorithm AS-36. Exact Confidence Limits for the Odds Ratio in a 2×2 Table," *Applied Statistics*, 20, 105–110.

Valz, P. D. and Thompson, M. E. (1994), "Exact Inference for Kendall's S and Spearman's Rho with Extensions to Fisher's Exact Test in $r \times c$ Contingency Tables," *Journal of Computational and Graphical Statistics*, 3(4), 459–472.

van Elteren, P. H. (1960), "On the Combination of Independent Two-Sample Tests of Wilcoxon," *Bulletin of the International Statistical Institute*, 37, 351–361.

Vollset, S. E., Hirji, K. F., and Elashoff, R. M. (1991), "Fast Computation of Exact Confidence Limits for the Common Odds Ratio in a Series of 2×2 Tables," *Journal of the American Statistical Association*, 86, 404–409.

Wilson, E. B. (1927), "Probable Inference, the Law of Succession, and Statistical Inference," *Journal of the American Statistical Association*, 22, 209–212.

Woolf, B. (1955), "On Estimating the Relationship Between Blood Group and Disease," *Annals of Human Genetics*, 19, 251–253.

Zelen, M. (1971), "The Analysis of Several 2×2 Contingency Tables," *Biometrika*, 58, 129–137.

Subject Index

Syntax Index

V

W

Z

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to **yourturn@sas.com**. Include the full title and page numbers (if applicable).
- If you have comments about the software, please send them to **suggest@sas.com**.