Cochran-Armitage Trend Test Using SAS

Hui Liu, Merck Research Labs, Merck & Co., Inc, Rahway, NJ

ABSTRACT
In clinical trials, a dose response study is often conducted to investigate the relationship between increasing dosage and the effect of the drug under study. The Cochran-Armitage trend test is commonly used to study the underlying trend. This paper summarizes various SAS® procedures, including PROC FREQ, PROC MULTTEST, and PROC LOGISTIC, which perform the trend test when the response is binary. Variations of Cochran-Armitage trend test, including asymptotic test, exact permutation test, monte carlo (permutation) resampling adjusted test, and bootstrap resampling adjusted test are discussed as well.

INTRODUCTION
In clinical trials, it is often of interest to investigate the relationship between the increasing dosage and the effect of the drug under study. Usually the dose levels tested are ordinal, and the effect of the drug is measured in binary. In this case, Cochran-Armitage trend test (Cochran, 1954; Armitage, 1955) is most frequently used to test for trend among binomial proportions.

Suppose the dose levels under tested are 1, 2, …, I, and the response is recorded as 1 or 0, which could represent success/failure, tumor/no tumor, rescue medication taken/not taken, etc. So the collected data can be summarized by a 2 × I table with ordered columns indicating dosage level \( i \), and two rows indicating the binary response.

<table>
<thead>
<tr>
<th>Dosage</th>
<th>1</th>
<th>2</th>
<th>( i )</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( n_{11} )</td>
<td>( n_{12} )</td>
<td>( n_{1i} )</td>
<td>( n_{1I} )</td>
</tr>
<tr>
<td>0</td>
<td>( n_{01} )</td>
<td>( n_{02} )</td>
<td>( n_{0i} )</td>
<td>( n_{0I} )</td>
</tr>
</tbody>
</table>

Here, the Null hypothesis (\( H_0 \)):
There is no linear trend in binomial proportions of response across increasing levels of dosage.

Alternative hypothesis (\( H_1 \)):
There is a linear trend in binomial proportions of response across increasing levels of dosage.

Agresti (1990) gave the test statistics for asymptotic Cochran-Armitage trend test. In column \( i \), let \( \pi_{1i} \) denote the real underlying probability of response 1, and let \( p_{1i} \) denote the observed sample proportion of response 1, \( i = 1, \ldots, I \). Let \( \{s_i\} \) be scores assigned to the columns. For the linear probability model

\[
\pi_{1i} = \alpha + \beta s_i,
\]

the hypothesis can be written into:

\[
H_0 : \pi_{11} = \pi_{12} = \ldots = \pi_{1I}
\]

versus the one-sided alternative hypothesis

\[
H_1 : \pi_{11} \leq \pi_{12} \leq \ldots \leq \pi_{1I}
\]

with at least one strict inequality.
The prediction equation under ordinary least squares fit is

\[ \hat{\pi}_{ij} = p_{1+} + b(s_i - \bar{s}), \]

where

\[ p_{1+} = \frac{n_{1+}}{n_{1+}}, \text{ and } \bar{s} = \frac{\sum n_{1i}s_i}{n}, \text{ and } b = \frac{\sum n_{1i}(p_{ij} - p_{1+})(s_i - \bar{s})}{\sum n_{1i}(s_i - \bar{s})^2}, \]

The test statistic for Cochran-Armitage trend test

\[ z^2 = \left( \frac{b^2}{p_{1+}p_{0+}} \right) \sum n_{1i}(s_i - \bar{s})^2, \]

where \( p_{0+} = 1 - p_{1+} \), has an asymptotic chi-squared distribution with df=1, or

\[ z = \frac{\sum n_{1i}(p_{ij} - p_{1+})(s_i - \bar{s})}{\sqrt{p_{1+}p_{0+}\sum n_{1i}(s_i - \bar{s})^2}}, \]

has an asymptotic normal distribution. Results of the Cochran-Armitage trend test are similar to those obtained by testing that the slope is zero in a linear logistic model.

Since binomial distribution can only take discrete values, continuity correction can be applied to better approximate the continuous distribution. With continuity correction \( c \), the upper-tailed \( p \)-value is computed from

\[ z = \frac{\sum n_{1i}(p_{ij} - p_{1+})(s_i - \bar{s}) - c}{\sqrt{p_{1+}p_{0+}\sum n_{1i}(s_i - \bar{s})^2}}. \]

The Cochran-Armitage trend test can be performed based on the asymptotic normality of its test statistic, as shown above, or based on an exact null distribution.

If we fix the margins of the contingency table and consider all the resulted possible permutations, the exact permutation \( p \)-value can be obtained by conditioning on the sum of the overall observed responses \( n_{1+} \) (Corcoran and Mehta, 2002). Let \( t_1, t_2, \ldots, t_M \) denote the \( M \) possible contingency tables with sum of responses \( n_{1+} \). The trend score calculated from each of these contingency tables is

\[ T(t_m | n_{1+}) = \sum n_{1i}^{(m)} s_{ii}, \]

where \( m = 1, 2, \ldots, M \).

Then the exact permutation \( p \)-value is evaluated as

\[ \sum_{m=1}^{M} p(t_m | n_{1+})I\{T(t_m | n_{1+}) \geq T(\text{observed})\}, \]
assuming the score is increasing with the increasing dosage, i.e., \( s_0 < s_1 < ... < s_I \). Here \( p(t_m \mid n_{1+}) \) is the conditional probability of observing table \( t_m \), and can be obtained by

\[
p(t_m \mid n_{1+}) = \prod_{i=1}^{I} \left( \frac{n_{+i}}{n_{i+}} \right)^{n_{++} \choose n_{1+}}
\]

In other words, the exact permutation \( p \)-value is the proportion of those contingency tables which have the Cochran-Armitage trend test score at least as extreme as the observed table, conditional on the sum of the response is the same as that for observed table.

For large sample size where the exact permutation is not feasible due to the restriction of the computation resource, permutation \( p \)-value can be obtained by resampling tables \( t_m \) with probability \( p(t_m \mid n_{1+}) \). Let \( N \) denote the total number of resampled tables, usually the number is large, then the permutation \( p \)-value estimate is the proportion of resampled tables with Cochran-Armitage trend test score at least as extreme as the observed table:

\[
\sum_{i=1}^{N} I\{T(t_m \mid n_{1+}) \geq T(\text{observed})\}
\]

\[
\frac{1}{N}
\]

Another resampling method is bootstrap. Bootstrap \( p \)-value is obtained by resampling tables with fixed columns’ sums \( n_{+i} \) with probability

\[
\prod_{i=1}^{I} \left( \frac{n_{+i}}{n_{i+}} \right)^{n_{++} \choose n_{1+}} \pi^{n_{++} \choose n_{1+}} (1 - \pi)^{n_{++} \choose n_{1+} - n_{1+}},
\]

where \( \pi \) can be replaced with the MLE estimate

\[
\hat{\pi} = \frac{n_{1+}}{n_{++}}.
\]

Then the bootstrap \( p \)-value estimate is the proportion of resampled tables with Cochran-Armitage trend test score at least as extreme as the observed table.

**ASYMPTOTIC METHOD**

Suppose a study is conducted to test the effect of a drug on 40 subjects. The subjects are randomized into four balanced groups receiving 0 mg, 1 mg, 2 mg and 3 mg of the drug, respectively. The results for each of the two responses are recorded for each subject, and the raw data can be created as follows:

data doseresp;
   input dose resp1 resp2 @@;
datalines;
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0
1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0
2 0 1 2 0 0 2 0 0 2 0 0 2 0 0 2 0
2 0 0 2 0 0 2 0 0 2 0 0 2 0 0 2 0
3 1 1 3 1 1 3 1 1 3 0 1 3 0

```r
data doseresp;
   input dose resp1 resp2 @@;
datalines;
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0
1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0
2 0 1 2 0 0 2 0 0 2 0 0 2 0 0 2 0
2 0 0 2 0 0 2 0 0 2 0 0 2 0 0 2 0
3 1 1 3 1 1 3 1 1 3 0 1 3 0
```
The dosage under tested has 4 levels (DOSE=0, 1, 2, 3). Within each dose level, two binary responses (RESP1, RESP2) are recorded for each of the ten subjects. The increasing incidence of 1s for higher dosages in the data suggests a possible relationship between dosage and drug effect. Cochran-Armitage trend test is used to test the possibility of the response increases linearly with drug level. Let's focus on response Resp1 first.

1) PROC FREQ
To perform the asymptotic Cochran-Armitage trend test, Base SAS® PROC FREQ procedure provides a simple way.

```sas
proc freq data=doseresp;
table resp1*dose / trend norow nocol nopercent scores=table;
run;
```

Here the TREND option is specified on the TABLE statement, which gives asymptotic Cochran-Armitage trend test statistics and p-values for one-sided and two-sided tests. Note the input table must be 2xC or Rx2 to do the trend test by this way. And as usual, this procedure gives 2-way table to summarize the input data. The results are shown as follows:

```
The FREQ Procedure
Table of resp1 by dose

resp1    dose
Frequency       0'       1'       2'       3'  Total
0'     10 '      9 '     10 '      7 '     36
1 '      0 '      1 '      0 '      3 '      4
Total          10       10       10       10       40

Statistics for Table of resp1 by dose
Cochran-Armitage Trend Test
Statistic (Z)       -1.8856
One-sided Pr < Z    0.0297
Two-sided Pr > |Z|   0.0593
```

From the statistic formula for asymptotic test shown above, scores for each dose level must be specified in order to calculate the test statistics. In this procedure, option SCORES=TABLE on the TABLE statement is the default score type, which uses the numeric dose levels from the input data as the scores (0, 1, 2, 3 in this example). Other choices for SCORES include RANK, RIDIT, MODRIDIT. They give the same results if the dose intervals are equal.

2) PROC MULTTEST
The same tests can be performed by using SAS/STAT® PROC MULTTEST procedure.

```sas
proc multtest data=doseresp notables;
class dose;
test ca(resp1 / binomial continuity=0 uppertailed);
contrast 'CA Linear Trend' 0 1 2 3;
run;
```
CA on TEST statement requests the Cochran-Armitage linear trend test for group comparisons. The response variable should be coded 1|0, where value 1 is for a success, and value 0 is for a failure. Binary coding with other values will not work. By default, a Z-score approximation is used to estimate the p-value. If we'd like to do the continuity correction for the Z-score approximation, the CONTINUITY option is available. Note by default, hypergeometric variance estimate is used for Cochran-Armitage test in asymptotic normal approximation. To match results from PROC FREQ, which uses the binomial variance estimate, we need to specify the BINOMIAL option on the TEST statement. By default, all tests are two-tailed. For one-sided test, we need to specify the direction by UPPERTAILED or LOWERTAILED options. In this procedure, scores are given in the CONTRAST statement. All of the scores need to be explicitly listed. Otherwise, default values of 0, 1, 2, ... will be used for scores. An important feature of the test is that the result is not changed if all the scores are added or multiplied by a common factor. Output from this procedure is shown as follows:

<table>
<thead>
<tr>
<th>The Multtest Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Information</td>
</tr>
<tr>
<td>Test for discrete variables</td>
</tr>
<tr>
<td>Z-score approximation used</td>
</tr>
<tr>
<td>Continuity correction</td>
</tr>
<tr>
<td>Tails for discrete tests</td>
</tr>
<tr>
<td>Strata weights</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>dose</td>
</tr>
<tr>
<td>Contrast 0 1 2 3</td>
</tr>
<tr>
<td>Trend 0 1 2 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>resp1</td>
</tr>
</tbody>
</table>

3) PROC LOGISTIC
Another procedure that can perform asymptotic Cochran-Armitage trend test is PROC LOGISTIC. The score test in the Testing Global Null Hypothesis: BETA=0 table is equivalent to the Cochran-Armitage trend test.

```
proc logistic data=doseresp descending;
  model resp1 = dose;
run;
```

Here we treat dose as a continuous variable in order to test the null hypothesis of beta=0. The p-value estimate in the output is the same result for the two-sided Cochran-Armitage trend test.

<table>
<thead>
<tr>
<th>Testing Global Null Hypothesis: BETA=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
</tr>
<tr>
<td>Score</td>
</tr>
<tr>
<td>Wald</td>
</tr>
</tbody>
</table>

All of the above three procedures give the same p-value for the asymptotic Cochran-Armitage trend test. The test supports the trend hypothesis since the p-value for the one-sided test is 0.0297, less than the well-accepted cutoff 0.05. The small p-value for the Cochran-Armitage test indicates that the probability of the Row 1 level (RESP1=0) decreases as DOSE increases or, equivalently, that the probability of the Row 2
level (RESP=1) increases as DOSE increases. The two-sided p-value tests against either an increasing or decreasing alternative. This is an appropriate hypothesis when the direction of the trend is unknown. The asymptotic p-value approximates the probability obtained from the exact permutation distribution, discussed in the following text.

**EXACT PERMUTATION METHOD**

For a small, sparse, or skewed sample, the asymptotic test might not be appropriate due to the invalid underlying normal distribution assumption. So the exact test is a better choice and gives accurate results.

1) **PROC FREQ**

PROC FREQ can get exact p-value for Cochran-Armitage trend test as well, by adding EXACT TREND statement:

```sql
proc freq data=doseresp;
  table resp1*dose / trend norow nocol nopercent;
  exact trend / maxtime=60;
run;
```

Since the exact test can take a long time, MAXTIME option on EXACT statement can be used to specify the maximum clock time (in seconds) that an exact p-value is computed. At that time the procedure stops regardless of the computation status. The results from output are shown here:

```
Exact Test
One-sided Pr <= Z    0.0507
Two-sided Pr >= |Z|   0.1014
```

2) **PROC MULTTEST**

The exact p-value can also be obtained from PROC MULTTEST procedure.

```sql
proc multtest data=doseresp notables;
  class dose;
  test ca(resp1 / permutation=5 uppertailed);
run;
```

Option PERMUTATION on the TEST statement specifies that exact permutation distributions are used when marginal success or failure totals are equal to or less than the number specified. In our example, exact permutation is used everywhere since the marginal success total is 4 for Resp1, which is less than 5. The output is shown here:

```
The Multtest Procedure

Model Information

Test for discrete variables       Cochran-Armitage
Exact permutation distribution used Everywhere
Tails for discrete tests          Upper-tailed
Strata weights                    None

p-Values

Variable  Contrast  Raw
resp1      Trend     0.0507
```

If we change PERMUTATION from 5 to 3, then for marginal totals greater than the specified number 3, PROC MULTTEST uses normal approximations to get p-values. Then the Cochran-Armitage trend test is implemented using a combination of both a Z-score approximation and an exact permutation distribution. The results are changed to:
3) PROC LOGISTIC
PROC LOGISTIC can generate exact p-values for a two-sided test by adding the EXACT statement. Actually this is the exact conditional scores test based on permutation or exact conditional distribution (Derr, 2000).

proc logistic data=doseresp descending exactonly
  exactoptions(maxtime=60 method=direct);
  model resp1=dose;
  exact dose;
run;

EXACTONLY option on the PROC LOGISTIC statement requests only the exact analyses. METHOD=DIRECT suboption in EXACTOPTIONS option, which is the default, directly builds the exact distribution.

### Conditional Exact Tests

<table>
<thead>
<tr>
<th>Effect</th>
<th>Test</th>
<th>Statistic</th>
<th>Exact</th>
<th>Mid</th>
</tr>
</thead>
<tbody>
<tr>
<td>dose</td>
<td>Score</td>
<td>3.4667</td>
<td>0.1014</td>
<td>0.0838</td>
</tr>
<tr>
<td></td>
<td>Probability</td>
<td>0.0353</td>
<td>0.1014</td>
<td>0.0838</td>
</tr>
</tbody>
</table>

The above three procedures give the same p-value for the exact Cochran-Armitage trend test. It differs from the p-value obtained from the asymptotic Cochran-Armitage trend test. For medium to large sample size, the exact test might not be appropriate since it uses many computational resources and requires a great amount of time. If we suspect that asymptotic approximations might not be sufficient, we can turn to permutation resampling estimation or bootstrap resampling estimation for the exact p-value estimate, as discussed next.

### RESAMPLING ADJUSTED METHOD
Instead of having a complete enumeration of all contingency tables under specific condition, sample tables can be drawn from the permutation or bootstrap distribution. The trend test can be tested through comparing the observed trend test score with an empirical distribution of values for the trend test statistic obtained by resampling. Permutation uses resampling without replacement, while bootstrap uses resampling with replacement.

1) PROC FREQ
We can request permutation estimation (also called Monte-Carlo estimation) in PROC FREQ by adding option MC on the EXACT statement.

 proc freq data=doseresp;
Here multiple options are specified on the EXACT statement. MC options requests Monte Carlo estimation of exact p-values. N specifies the number of samples for MC estimation, 10,000 is the default value. SEED specifies the initial seed for random number generation for MC estimation. The results are

Monte Carlo Estimates for the Exact Test

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>One-sided Pr &lt;= Z Estimate</td>
<td>0.0507</td>
<td></td>
</tr>
<tr>
<td>99% Lower Conf Limit</td>
<td>0.0504</td>
<td></td>
</tr>
<tr>
<td>99% Upper Conf Limit</td>
<td>0.0510</td>
<td></td>
</tr>
<tr>
<td>Two-sided Pr &gt;=</td>
<td>Z</td>
<td>Estimate</td>
</tr>
<tr>
<td>99% Lower Conf Limit</td>
<td>0.1012</td>
<td></td>
</tr>
<tr>
<td>99% Upper Conf Limit</td>
<td>0.1019</td>
<td></td>
</tr>
<tr>
<td>Number of Samples</td>
<td>5000000</td>
<td></td>
</tr>
<tr>
<td>Initial Seed</td>
<td>123456</td>
<td></td>
</tr>
</tbody>
</table>

Monte-Carlo estimates are close enough to the exact p-values if the number of samples is sufficiently large.

2) PROC MULTTEST

PROC MULTTEST can also calculate adjusted p-values from a trend test by using permutation or bootstrap resampling method.

The PERMUTATION option on PROC MULTTEST statement requests permutation resampling adjusted p-value. The NSAMPLE option specifies the number of samples for use (default value is 20,000). A large number of samples is usually used for accuracy, but this results in long execution times, particularly with large data sets.

```
proc multtest data=doseresp notables
   permutation nsample=5000000 seed=123456;
   class dose;
   test ca(resp1 / permutation=5 uppertailed);
run;
```

With sufficiently large number of samples, permutation resampling gives the estimate of p-value very close to the exact permutation p-value.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Contrast</th>
<th>Raw</th>
<th>Permutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>resp1</td>
<td>Trend</td>
<td>0.0507</td>
<td>0.0506</td>
</tr>
</tbody>
</table>

The Multtest Procedure

Model Information

<table>
<thead>
<tr>
<th>Test for discrete variables</th>
<th>Cochran-Armitage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact permutation distribution used</td>
<td>Everywhere</td>
</tr>
<tr>
<td>Tails for discrete tests</td>
<td>Upper-tailed</td>
</tr>
<tr>
<td>Strata weights</td>
<td>None</td>
</tr>
<tr>
<td>P-value adjustment</td>
<td>Permutation</td>
</tr>
<tr>
<td>Number of resamples</td>
<td>5000000</td>
</tr>
<tr>
<td>Seed</td>
<td>123456</td>
</tr>
</tbody>
</table>

p-Values

<table>
<thead>
<tr>
<th>Variable</th>
<th>Contrast</th>
<th>Raw</th>
<th>Permutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>resp1</td>
<td>Trend</td>
<td>0.0507</td>
<td>0.0506</td>
</tr>
</tbody>
</table>
For bootstrap adjusted \( p \)-values, the \texttt{BOOTSTRAP} option is available for use on \texttt{PROC MULTTEST} statement. The options used here are similar to those used in the permutation method. If we specify \texttt{PERMUTATION=} suboption on \texttt{TEST} statement, the procedure takes a long time since the exact permutation distribution is recomputed for each bootstrap sample. So this suboption is preferably used when permutation is used in the base test.

\begin{verbatim}
proc multtest data=doseresp notables
   bootstrap nsample=5000000 seed=123456;
   class dose;
   test ca(resp1 / binomial uppertailed);
run;
\end{verbatim}

Bootstrap resampling adjusted \( p \)-value is different from the permutation resampling adjusted \( p \)-value. The accuracy of bootstrap \( p \)-value depends on how well the MLE estimate of \( \pi \), the probability of response in population, instead of the size of resample (Corcoran and Mehta, 2002).

\begin{verbatim}
Bootstrap resampling adjusted p-value is different from the permutation resampling adjusted p-value. The accuracy of bootstrap p-value depends on how well the MLE estimate of \( \pi \), the probability of response in population, instead of the size of resample (Corcoran and Mehta, 2002).
\end{verbatim}

\begin{verbatim}
 proc multtest data=doseresp notables
   permutation nsample=5000000 seed=123456;
   class dose;
   test ca(resp1 resp2 / permutation=6 uppertailed);
run;
\end{verbatim}

\textbf{RESAMPLING METHOD FOR MULTIPLE TESTS}

We could end up measuring two or more responses across the ordinal dose level, just like in our example data. If we try to do the Cochran-Armitage trend test for all the responses, we will get an increased probability of declaring false significances. Given this, we have to address the multiplicity problem. A big advantage of using \texttt{PROC MULTTEST} is doing multiple tests and concurrently dealing with the multiplicity issue by controlling the overall type I error. \texttt{PROC MULTTEST} can adjust the \( p \)-values from multiple Cochran-Armitage trend tests. This procedure controls the family-wise error rate at or below the type-I error level, by incorporating the correlational and distributional characteristics.

The following statements perform a typical analysis by using permutation:

\begin{verbatim}
 proc multtest data=doseresp notables
   permutation nsample=5000000 seed=123456;
   class dose;
   test ca(resp1 resp2 / permutation=6 uppertailed);
run;
\end{verbatim}

\begin{verbatim}
 p-Values
<table>
<thead>
<tr>
<th>Variable</th>
<th>Contrast</th>
<th>Raw</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>resp1</td>
<td>Trend</td>
<td>0.0297</td>
<td>0.0354</td>
</tr>
<tr>
<td>resp2</td>
<td>Trend</td>
<td>0.0382</td>
<td>0.0492</td>
</tr>
</tbody>
</table>
\end{verbatim}
Bootstrap adjusted method for multiple tests is shown as follows:

```
proc multtest data=doseresp notables
  bootstrap nsample=5000000 seed=123456;
  class dose;
  test ca(resp1 resp2 / binomial uppertailed);
run;
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Contrast</th>
<th>Raw</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>resp1</td>
<td>Trend</td>
<td>0.0297</td>
<td>0.0620</td>
</tr>
<tr>
<td>resp2</td>
<td>Trend</td>
<td>0.0238</td>
<td>0.0337</td>
</tr>
</tbody>
</table>

The Raw column lists the p-values for the Cochran-Armitage test on the original data, and the Permutation and Bootstrap columns provide the permutation/bootstrap adjustment of the raw p-values, respectively. Failure to account for the multiple testing issues can certainly lead to misleading inferences for the data.

**CONCLUSION**

Cochran-Armitage trend test is a frequently used test in dose-response studies. In this paper, multiple approaches to do the test with binary response are shown here and the underlying theories are discussed as well. For large samples where normal distribution assumption is valid, asymptotic Cochran-Armitage trend test is handy and appropriate. All the three SAS procedures discussed here, PROC FREQ, PROC MULTTEST, and PROC LOGISTIC can implement this test, while giving the same p-values. For small sample sizes, we need to turn to the exact permutation Cochran-Armitage trend test since the normal distribution assumption does not hold. All the three SAS procedures discussed can do the exact test. For medium sample sizes, where we suspect the validity of the normal distribution assumption, and the exact permutation might take a lot of computation resources, we can turn to the resampling method by either permutation (without replacement) or bootstrap (with replacement). PROC MULTTEST can handle multiple Cochran-Armitage trend tests by controlling the type-I error level, through the permutation or bootstrap resampling method.

If the data set has a variable that contains the frequency of occurrence of each observation, WEIGHT statement in PROC FREQ, FREQ statement in PROC MULTTEST and PROC LOGISTIC can do the test by treating each observation as if it appears $n$ times.

In this paper, we choose the scores 0, 1, 2, 3, … from the default settings or the numeric dose levels. But we have choices in selecting these scores, especially with procedure PROC MULTTEST. Scores have to be meaningful and reflecting the degree of difference among the dose levels. The test result is invariant to the location and scale of the scores.

Stratum is not discussed in this paper. But it is very important to the trend test. If the data contains the stratum variable, we need to include it in the trend test procedures before concluding the final results.

**REFERENCES**

SAS OnlineDoc® 9.1.3, SAS Institute Inc.

**ACKNOWLEDGEMENTS**

The author would like to thank Dr. Kaifeng Lu, Dr. Haiyuan Zhu, and Elaine Czamecki for their careful review and helpful comments. The author also thanks the management team for their review of this paper and help to go through the company publication process.
CONTACT INFORMATION
Your comments and questions are valued and encouraged. Contact the author at:

Hui Liu
Merck & Co., Inc.
RY34-A320
P.O. Box 2000
Rahway, NJ 07065
Phone: (732) 594-5732
Fax: (732) 594-6075
Email: hui_liu@merck.com

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

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