BIOS 6244 Analysis of Categorical Data
October 26, 2005 Supplement

Exact Inference for Conditional Associations (Sec. 3.3)

As in the case of performing statistical inference for 2x2 tables, we must be concerned about the accuracy of χ^2 approximations to the distributions of the test statistics when performing inference for conditional odds ratios. The accuracy of the χ^2 approximations depends more on the size of the marginal totals than on the individual cell counts in the partial tables. If one or more of the partial tables has small marginal totals, the χ^2 approximations may be inaccurate and alternative methods must be used. Exact procedures inference for conditional odds ratios have been developed, and should be used when available. In this section, we briefly describe these exact procedures and illustrate them with an example.

Exact Test of Conditional Independence for 2x2xK Tables (Sec. 3.3.1)

The CMH test statistic focuses on the observed cell counts $\{n_{11k}\}$ in Row 1, Column 1 of each partial table. The exact version of the CMH test uses $\sum_k n_{11k}$ and its exact distribution in much the same way that Fisher's exact test uses n_{11} and its exact distribution. Conditional on the marginal totals in the k'th partial table, the exact distribution of n_{11k} is hypergeometric. The exact distribution of the CMH test statistic is derived using the exact distribution of $\sum_k n_{11k}$.

Recall that the null hypothesis of conditional independence is equivalent to the null hypothesis that $OR_{XY(k)} = 1$ for all k = 1, 2, .., K. A "positive" conditional association corresponds to the one-sided alternative hypothesis that $OR_{XY(k)} > 1$ for k = 1, 2, .., K. To perform this upper-tailed test,we need an upper-tailed p-value. This p-value is equal to the probability that $\sum_k n_{11k}$ is at least as large as the observed value, given the marginal totals in each partial table.

Similarly, for the 1-sided alternative $OR_{XY(k)} < 1$ for all k = 1, 2, .., K, we must perform a lower-tailed test, and the p-value is equal to the probability that $\sum_k n_{11k}$ is no larger than the observed value, given the marginal totals in each table. For a 2-sided alternative, we accumulate the exact probability for $\sum_k n_{11k}$ across all partial tables in the reference set that are no more likely than the one observed.

Example (alcohol vs. esophageal cancer)

Consider the following data obtained from a case-control study examining the association between alcohol exposure and esophageal cancer in younger patients, stratified by age.

Age Group	Alcohol Exposure		No Alcohol Exposure	
	Case	Control	Case	Control
25-34	1	9	1	106
35-44	2	26	2	164

Consider the partial tables for the 25-34 and 35-44 age groups. We have $n_{111} = 1$ and $n_{112} = 2$. Given that the column totals are fixed (as in all case-control studies), n_{111} can take on the possible values of 0, 1, 2 (since the total number of cases in layer 1 is 2). Similarly, n_{112} can take on the possible values of 0, 1, 2, 3, 4. Then the test statistic for the exact version of the CMH test $\sum_k n_{11k}$ can take on integer values from 0 to 6. The observed value is 3 and any values of $\sum_k n_{11k}$ larger than 3 are favorable to the alternative hypothesis. Thus, the upper-tailed p-value in this case is given by P(3) + P(4) + P(5) + P(6), where each probability is calculated using the exact distribution of $\sum_k n_{11k}$. These calculations are extremely tedious to perform by hand, so we resort to using statistical software. Unfortunately, neither SAS nor SPSS can perform the exact version of the CMH test. Using StatXact, the exact 2-tailed p-value is .029, indicating that the conditional OR's are not all equal to 1. We will consider the use of SAS in performing CMH and related analyses below.

An interesting application of the exact version of the CMH test to promotion discrimination based on race is given in Sec. 3.3.2 (Table 3.4), p. 65 of our text.

<u>Exact Test of Homogeneity of Odds Ratios (Sec. 3.3.4)</u>

The Breslow-Day test discussed in Sec. 3.2.4 is only an approximate test. For a small total sample size n or a small sample size in one or more of the partial tables, the χ^2 approximation may not be valid. An exact test of homogeneity, called *Zelen's test*, is available and should be used whenever possible, especially under the above circumstances. In a manner similar to the method used to find the exact distribution of Fisher's exact test, the exact distribution of Zelen's test statistic is calculated using the reference set of all 2x2xK tables that have the same 2-way marginal totals as the observed table. The exact p-value is then the sum of the probabilities for all 2x2xK tables in the reference set that are no more likely than the observed table.

Exact Hypothesis Test and CI for the Common OR (Sec. 3.3.3)

An exact method is available for testing the null hypothesis H_0: $OR_{XY} = 1$. To find an exact 95% CI(OR_{XY}), Agresti recommends the approach that was used to find an exact 95% CI(OR) in Section 2.6.4; namely, to include in the exact 95% CI(OR_{XY}) all null values of OR_{XY} that would not be rejected in an hypothesis test using $\alpha = .05$. As we saw in Sec. 2.6, exact hypothesis tests and exact CI's based on discrete test statistics tend to be conservative. As a result, Agresti recommends that the mid-p approach be used when performing the exact test of H_0: $OR_{XY} = 1$ and when finding an exact 95% CI(OR_{XY}). The mid-p adjustment is available in StatXact, but not in SAS or SPSS.

Use of SAS in Performing CMH Analyses

Reading Assignment: Stokes et al., Chapter 3, pp. 43-66

The following SAS code (available on the course website) creates the 2x2x2 table for the data given above on a possible association between alcohol exposure and esophageal cancer and performs the Cochran-Mantel-Haenszel Analyses covered in Sec. 3.2 in our text.

```
data esophageal;
input alcohol case agegp count @@;
cards;
1 1 1 1  1 2 1 9   2 1 1 1  2 2 1 106
1 1 2 2  1 2 2 26  2 1 2 2  2 2 2 164
;

proc freq order = data; weight count;
   tables  agegp * alcohol * case / nocol nopct cmh bdt relrisk chisq;
   exact or chisq ;
   title 'Alcohol vs. Esophageal Cancer Example';
   title2 'Cochran-Mantel-Haenszel Analyses';
   run;
```

The relevant output produced by this code is as follows:

```
Alcohol vs. Esophageal Cancer Example
  Cochran-Mantel-Haenszel Analyses

      The FREQ Procedure

Statistics for Table 1 of alcohol by case
      Controlling for agegp=1

      Pearson Chi-Square Test
------------------------------
Chi-Square                  4.4732
DF                               1
Asymptotic Pr >  ChiSq      0.0344
Exact       Pr >= ChiSq     0.1643
```

(This is the exact χ^2 test in layer 1.)

(Output continued on next page.)

```
   Odds Ratio (Case-Control Study)
-----------------------------------
Odds Ratio                  11.7778
```
(This is the estimated OR in layer 1.)

```
Asymptotic Conf Limits
95% Lower Conf Limit         0.6785
95% Upper Conf Limit       204.4517

Exact Conf Limits
95% Lower Conf Limit         0.1351
95% Upper Conf Limit       932.6034
```
[This is the exact 95% CI(OR)in layer 1.]

```
Alcohol vs. Esophageal Cancer Example
  Cochran-Mantel-Haenszel Analyses

         The FREQ Procedure

Statistics for Table 2 of alcohol by case
      Controlling for agegp=2

      Pearson Chi-Square Test
----------------------------------
Chi-Square                  4.1835
DF                               1
Asymptotic Pr >  ChiSq      0.0408
Exact      Pr >= ChiSq      0.1003
```
(This is the exact χ^2 test in layer 2.)

```
   Odds Ratio (Case-Control Study)
-----------------------------------
Odds Ratio                   6.3077
```
(This is the estimated OR in layer 2.)

```
Asymptotic Conf Limits
95% Lower Conf Limit         0.8510
95% Upper Conf Limit        46.7552

Exact Conf Limits
95% Lower Conf Limit         0.4329
95% Upper Conf Limit        89.1502
```
[This is the exact 95% CI(OR)in layer 2.]

```
Alcohol vs. Esophageal Cancer Example
  Cochran-Mantel-Haenszel Analyses

        The FREQ Procedure

Summary Statistics for alcohol by case
       Controlling for agegp

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

Statistic    Alternative Hypothesis    DF       Value      Prob
---------------------------------------------------------------
    1        Nonzero Correlation        1      7.9064    0.0049
```
(This is the approx. Cochran-Mantel-Haenszel test.)

(Computer output continued on next page.)

```
Estimates of the Common Relative Risk (Row1/Row2)

Type of Study      Method                     Value      95% Confidence Limits
--------------------------------------------------------------------------
Case-Control       Mantel-Haenszel           7.5275        1.4509       39.0535
  (Odds Ratio)     Logit                     7.7512        1.5041       39.9439
                                                       (The highlighted values
  Breslow-Day-Tarone Test for                          are an approx. 95% CI for
Homogeneity of the Odds Ratios                         the common OR across
------------------------------                         layers.)
Chi-Square                  0.1247
DF                               1  (This is the Breslow-Day test for homogeneous
Pr > ChiSq                  0.7240   association, with the Tarone adjustment.)
```

Note that the OR's differ somewhat between the 2 layers (11.78 vs. 6.31) and that neither one is statistically significant. An "eyeball" comparison of the 2 OR's suggests that there may be interaction between alcohol exposure and age; that is, that the association between alcohol and esophageal cancer differs between age groups. However, this was not confirmed by the Breslow-Day test. It *is* appropriate to proceed with the Mantel-Haenszel estimation of the common OR since the direction of the partial OR's is the same in the 2 layers *and* the null hypothesis of homogeneous association is not rejected.

It is instructive to compare the M-H estimate of the OR (that is, the estimated OR after adjusting for the possible confounding effects of age) with the results obtained if we ignore age. PROC FREQ yields the following results based on the marginal table:

OR = 7.71
Exact 95% CI(OR) = (0.98, 59.17)

Note that the M-H estimate of the OR does not differ very much from the OR based on the marginal table (7.53 vs. 7.71). However, there is a rather extreme difference between the 2 CI(OR). This illustrates the importance of incorporating the effects of confounding variables whenever possible.

It is also instructive to compare the results for the approximate and exact versions of the various procedures:

Procedure	Approximate Result	Exact Result	Mid p Result
CMH Test	.005	.029	---
Test of Homogeneous Association	.724	1.000	--
Test that Common OR = 1	.005	.054	.030
CI for Common OR	(1.45, 39.05)	(0.97, 59.26)	(1.25, 45.92)

Note that using exact method yields results for the common OR that are not consistent with the "clinical" significance of the findings. An estimated OR as large as 7.5 typically has serious practical implications, yet the exact test and CI indicate that the common OR is no different from 1. Applying the mid p adjustment yields a p-value and CI for the common OR that make more sense in the context of the applied problem. Note the similarities between the approximate results for the common OR and those based on the mid p adjustment.