4.6.3 Zelen's Test

A computationally simple test of homogeneity was proposed by Zelen (1971). Let X_A^2 designate a test of the hypothesis H_{0A} of no partial association (or just association) for the average measure of association among the K strata, such as the Cochran-Mantel-Haenszel test on 1 *df* of the common odds ratio. Since the omnibus null hypothesis can be partitioned as shown in (4.55), Zelen (1971) proposed that the test of homogeneity

$$X_{H,Z}^2 = X_O^2 - X_A^2 \qquad (4.69)$$

be obtained as the difference between the omnibus chi-square test X_O^2 on K *df*, and the test of association X_A^2 on 1 *df*. For the ulcer clinical trial example using the conditional Mantel-Haenszel test yields $X_{H,Z}^2 = 7.4648 - 3.00452 = 4.46$, which is slightly less than the Cochran test value $X_{H,C}^2 = 4.58$.

Mantel, Brown, and Byar (1977) and Halperin, Ware, Byar, et al. (1977) criticize this test and present examples that show that this simple test may perform poorly in some situations. The problem is that an optimal test of the null hypothesis of homogeneity H_{0H} should use the variances estimated under that hypothesis. Thus both the contrast test X_H^2 in (4.64) and Cochran's test $X_{H,C}^2$ in (4.68) use the variances $\widehat{\Sigma}_{\widehat{\theta}}$ estimated under the general alternative hypothesis H_{1O} in (4.49) that some of the $\{\theta_j\}$, if not all, differ from the null value θ_0. However, both X_A^2 and X_O^2 use the variances defined under the general null hypothesis H_0, Σ_0, so that equality does not hold, that is, $X_O^2 \neq X_H^2 + X_A^2$. In general, therefore, this test should be avoided.

4.6.4 Breslow-Day Test for Odds Ratios

Breslow and Day (1980) also suggested a test of homogeneity for odds ratios for use with a Mantel-Haenszel test that is based on the Mantel-Haenszel estimate of the common odds ratio $\widehat{OR}_{MH}$ in (4.14). In the jth stratum, given the margins for that 2×2 table $(m_{1j}, m_{2j}, n_{1j}, n_{2j})$ then the expectation of the index frequency a_j under the hypothesis of homogeneity $OR_j = OR$ can be estimated as

$$\widehat{E}(a_j|\widehat{OR}_{MH}) = \widetilde{a}_j \text{ such that } OR_j = \widehat{OR}_{MH}. \qquad (4.70)$$

This expected frequency is the solution to

$$\frac{(\widetilde{a}_j)(n_{2j} - m_{1j} + \widetilde{a}_j)}{(m_{1j} - \widetilde{a}_j)(n_{1j} - \widetilde{a}_j)} = \widehat{OR}_{MH} \qquad (4.71)$$

(see also Problem 2.7.2). Solving for $\widetilde{a}_j$ yields

$$m_{1j}n_{1j}\widehat{OR}_{MH} = \qquad (4.72)$$
$$\widetilde{a}_j \left[n_{2j} - m_{1j} + \widehat{OR}_{MH}(n_{1j} + m_{1j})\right] + \widetilde{a}_j^2 \left(1 - \widehat{OR}_{MH}\right),$$

which is a quadratic function in $\widetilde{a}_j$. The root such that $0 < \widetilde{a}_j \le \min(n_{1j}, m_{1j})$ yields the desired estimate. Given the margins of the table $(n_{1j}, n_{2j}, m_{1j}, m_{2j})$ the expected values of the other cells of the table are obtained by subtraction, such as $\widetilde{b}_j = m_{1j} - \widetilde{a}_j$.

Then the Breslow-Day test of homogeneity of odds ratios is a Pearson contingency test of the form

$$X^2_{H,BD} = \sum_{j=1}^{K} \left[\frac{(a_j - \widetilde{a}_j)^2}{\widetilde{a}_j} + \frac{(b_j - \widetilde{b}_j)^2}{\widetilde{b}_j} + \frac{(c_j - \widetilde{c}_j)^2}{\widetilde{c}_j} + \frac{(d_j - \widetilde{d}_j)^2}{\widetilde{d}_j} \right]. \quad (4.73)$$

Since the term in the numerator is the same for each cell of the jth stratum, for example, $(b_j - \widetilde{b}_j)^2 = (a_j - \widetilde{a}_j)^2$, this statistic can be expressed as

$$X^2_{H,BD} = \sum_{j=1}^{K} \left[\frac{(a_j - \widetilde{a}_j)^2}{\widehat{V}(a_j|\widehat{OR}_{MH})} \right] \quad (4.74)$$

where

$$\widehat{V}(a_j|\widehat{OR}_{MH}) = \left[\frac{1}{\widetilde{a}_j} + \frac{1}{\widetilde{b}_j} + \frac{1}{\widetilde{c}_j} + \frac{1}{\widetilde{d}_j} \right]^{-1}. \quad (4.75)$$

This test for homogeneity of odds ratios is used in SAS PROC FREQ as part of the Cochran-Mantel-Haenszel analysis with the CMH option. For the data from Example 4.1, this test yields the value $X^2_{H,BD} = 4.626$ on 2 *df* with $p \le 0.099$ and for Example 4.6, this test value is $X^2_{H,BD} = 1.324$ on 3 *df* with $p \le 0.7234$. In both cases, the test value is slightly larger than the Cochran test value, the *P*-values smaller.

Breslow and Day (1980) suggested that $X^2_{H,BD}$ is distributed asymptotically as χ^2_{K-1} on $K - 1$ *df.* Tarone (1985) showed that this would be the case if a fully efficient estimate of the common odds ratio, such as the *MLE*, were used as the basis for the test. Since the Mantel-Haenszel estimate is not fully efficient, then $X^2_{H,BD}$ is stochastically larger than a variate distributed as χ^2_{K-1}. Tarone also showed that a corrected test can be obtained as

$$X^2_{H,BD,T} = X^2_{H,BD} - \frac{\left(\sum_{j=1}^{K} a_j - \sum_{j=1}^{K} \widetilde{a}_j\right)^2}{\sum_{j=1}^{K} \widehat{V}(a_j|\widehat{OR}_{MH})} \quad (4.76)$$

which is asymptotically distributed as χ^2_{K-1} on $K - 1$ *df.* Breslow (1996) recommends that in general the corrected test should be preferred to the original Breslow-Day test, but also points out that the correction term is often negligible. For the data in Example 4.1, the corrected test value $X^2_{H,BD,T} = 4.625$ on 2 *df* with $p \le 0.100$; and for Example 4.6 $X^2_{H,BD,T} = 1.3236$ on 3 *df* with $p \le 0.7235$; in both cases, nearly identical to the original Breslow-Day test.