6.3 Some applications involving binary data:

Combination of several 2×2 tables:

Suppose there are m centres participating in the clinical trial. Then, the following tables can be obtained, $j = 1, \cdots, m$,

	Success	Failure	
Treatment	y_{1j}	$n_{1j} - y_{1j}$	n_{1j}
Control	y_{2j}	$n_{2j} - y_{2j}$	n_{2j}
	s_{1j}	s_{2j}	n_j

We may consider $Y_{1j} \sim B(n_{1j}, \pi_{1j})$ and $Y_{2j} \sim B(n_{2j}, \pi_{2j})$ with

$$logit(\pi_{1j}) = \lambda_j + \Delta,$$

$$logit(\pi_{2j}) = \lambda_j.$$

The main difficulty based on the full likelihood is that it contains $m+1$ parameters $\lambda_1, \lambda_2, \cdots, \lambda_m, \Delta$, to be estimated. In such circumstances, maximum likelihood needs not be consistent or efficient for large m. However, if we used the conditional likelihood based on $Y_{1j} | Y_{1j} + Y_{2j} = s_{1j}$, the likelihood then depends on only one parameter Δ. The conditional log-likelihood for Δ is

$$l_c(\Delta) \propto \sum_{j=1}^{m} log\left[\frac{\varphi^{y_{1j}}}{P_0(\varphi)}\right] = \sum_{j=1}^{m} \{y_{1j} \cdot log(\varphi) - log[P_0(\varphi)]\}$$

$$= \sum_{j=1}^{m} \{y_{1j}\Delta - log[P_0(exp(\Delta))]\}$$

where

$$\varphi = \frac{\left(\frac{\pi_{1j}}{1-\pi_{1j}}\right)}{\left(\frac{\pi_{2j}}{1-\pi_{2j}}\right)},$$

$$\Delta = logit(\pi_{1j}) - logit(\pi_{2j}) = log\left(\frac{\pi_{1j}}{1-\pi_{1j}}\right) - log\left(\frac{\pi_{2j}}{1-\pi_{2j}}\right) = log(\varphi),$$

and

$$P_0(\varphi) = \sum_{k=a}^{b} \binom{n_{1j}}{k} \binom{n_{2j}}{s_{1j}-k} \varphi^k.$$

Thus, the maximum conditional likelihood estimate is the solution of

$$\frac{dl_c(\Delta)}{d\Delta} = 0 \cdot$$

In addition, provided the total conditional Fisher information is sufficiently large,

standard large-sample likelihood theory applies to the conditional likelihood. To test $H_0 : \Delta = 0$, we can use the score statistic under H_0:

$$U = \left[\frac{dl_c(\Delta)}{d\Delta}\right]_{\Delta=0} = \sum_{j=1}^{m}\left[y_{1j} - \frac{n_{1j}(y_{1j}+y_{2j})}{n_j}\right] = \sum_{j=1}^{m}\left(y_{1j} - \frac{n_{1j}s_{1j}}{n_j}\right),$$

where the conditional expectation Y_{1j} is

$$\frac{n_{1j}}{n_j} \cdot s_{1j}.$$

In addition,

$$Var(U) = \sum_{j=1}^{m}\frac{n_{1j}n_{2j}s_{1j}(n_j - s_{1j})}{n_j^2(n_j - 1)}.$$

.

Then, the test statistic is

$$Z^- = \frac{U - {}^1\!/_2}{\sqrt{Var(U)}}$$

for one-sided hypothesis. The test is known as the Mantel-Haenszel test.

Note:

There is the probability that Δ varies from centre to centre. That is, Δ might not be constant over different centres. Such interactions require careful investigation and detailed plausible explanation.