

## 2 × 2 분할표를 이용한 조건부 독립성 검정 \*

김지현<sup>1)</sup> 임현선<sup>2)</sup>

### 요약

역학연구에서 두 수준을 갖는 위험인자 X와 이항 반응변수 Y의 관계에 관심을 갖는 경우가 많다. 이 때 두 변수의 상관관계에 영향을 미칠 수 있는 인자 Z의 값을 제어함에 따라 X와 Y의 상관관계가 여전히 존재하는지를, 즉 X와 Y의 조건부 독립성을 검정할 필요가 있다. 관측값의 수가 많지 않을 때, X와 Y의 조건부 독립성 검정을 위해 Mantel-Haenszel 검정이 널리 사용되고 있다. 하지만 X와 Y의 상관관계가 Z의 수준에 따라 그 방향까지 변할 경우 이 검정은 낮은 검정력을 갖는다. 본 연구에서는 이 경우에 높은 검정력을 갖는 대안 검정통계량을 제안한다. 대안 검정통계량의 분포에 대해 알아보고 모의실험을 통해 Mantel-Haenszel 검정과 비교해 본다.

### 1. 서론

범주형 자료의 분석에서 조건부 독립성(conditional independence)을 검정하는 작업은 매우 중요하다. 어떤 역학연구(epidemiological study)에서 범주형 위험인자 X와 병의 발병 여부 Y의 관련성에 관심이 있다고 하자. 그래서 X와 Y에 관한 분할표를 작성하고 독립성 검정에서 유의한 결론을 얻었다고 하자. 이 때 이 관련성은 여기에 영향을 미칠 수 있는 다른 인자(배후인자) Z의 효과가 중첩됨(confounded)으로 인해서 생긴 허위관련성(spurious association)일 수 있다. 따라서 X와 Y에 관한 분할표를 배후인자의 값에 따라 층화하여, 동일한 배후인자의 값에서 X와 Y의 관련성이 여전히 존재하는가를 살펴보아야 한다. 즉, Z가 주어졌을 때 X와 Y의 조건부 독립성 검정을 실시하여야 한다. 본 연구에서는 조건부 독립성을 검정하는 데에 쓰이는 Cochran-Mantel-Haenszel 통계량의 문제점을 지적하고 대안 통계량을 제안하고자 한다.

Mantel-Haenszel (1959)은 배후인자(제어변수(control variable) 또는 장애변수(nuisance variable)라 부르기도 함)의 효과를 고려할 때의 두 그룹의 이항반응변수에 대한 효과를 비교하는 검정을 제시하였다. 범주형 변수 X, Y, Z에서, 제어변수 Z의 수준은 K개이고 X와 Y의 수준은 각각 2개씩일 때, 자료를 K개의 2 × 2 분할표로 나타낼 수 있다. 이 때 제어변수 Z의 수준수 K는 한 배후인자의 수준 수일 수도 있고, 여러 배후인자들의 가능한 모든 수준 조합의 수일 수도 있다. X의 i 수준, Y의 j 수준, Z의 k 수준의 칸의 도수를  $n_{ijk}$ 로 표시하고, 칸의 확률을  $\pi_{ijk}$ 로 표시하기로 하자.  $n_{i+k} = \sum_{j=1}^2 n_{ijk}$ ,  $n_{+jk} = \sum_{i=1}^2 n_{ijk}$ ,  $n_{++k} = \sum_i \sum_j n_{ijk}$ ,  $\pi_{i+k} = \sum_{j=1}^2 \pi_{ijk}$ ,  $\pi_{+jk} = \sum_{i=1}^2 \pi_{ijk}$ ,  $\pi_{++k} = \sum_i \sum_j \pi_{ijk}$ 라고 정의한다.  $n_{+1k}$ ,  $n_{+2k}$ ,  $n_{1+k}$ ,  $n_{2+k}$ 가 주어졌을 때  $n_{11k}$ 는 초기하분포를 따른다. 따라서

\* 본 연구는 1997년 숭실대학교 교내연구비의 지원을 받아 수행되었음.

1) (156-743) 서울시 동작구 상도5동, 숭실대학교 통계학과, 부교수

2) (110-791) 서울시 종로구 동숭동 169, 한국방송대학교 응용통계학과, 조교

조건부 독립(모든  $i, j, k$ 에 대하여  $\pi_{ij|k} = \pi_{i+|k}\pi_{+j|k}$ )이라는 가정 하에  $n_{11k}$ 의 평균과 분산은 각각

$$m_{11k} = E(n_{11k}) = \frac{n_{1+k}n_{+1k}}{n_{++k}} \quad (1.1)$$

$$V(n_{11k}) = \frac{n_{1+k}n_{2+k}n_{+1k}n_{+2k}}{n_{++k}^2(n_{++k} - 1)} \quad (1.2)$$

이다. 그리고 서로 다른 층의  $n_{11k}$ 들은 독립된 표본에서 관측한 값이므로 독립이다. 따라서  $\sum n_{11k}$ 의 평균과 분산은 각각  $\sum m_{11k}$ ,  $\sum V(n_{11k})$ 이다. Mantel과 Haenszel (1959)은 조건부 독립성의 검정통계량으로서

$$\frac{(|\sum n_{11k} - \sum m_{11k}| - \frac{1}{2})^2}{\sum V(n_{11k})}$$

을 제안하였다. 한편 Cochran(1954)은 연속성 수정항인  $\frac{1}{2}$ 을 쓰지 않고 다른 형태의 분산

$$\frac{n_{1+k}n_{2+k}n_{+1k}n_{+2k}}{n_{++k}^2}$$

을 고려한 통계량을 제안하였는데 두 통계량은 점근적으로 동등하다. Mantel-Haenszel 통계량에서 연속성 수정을 고려하지 않는 통계량

$$M^2 = \frac{(\sum n_{11k} - \sum m_{11k})^2}{\sum V(n_{11k})} = \frac{[\sum (n_{11k} - m_{11k})]^2}{\sum V(n_{11k})} \quad (1.3)$$

을 Cochran-Mantel-Haenszel(CMH) 통계량으로 정의하기로 한다. 조건부 독립성이 만족된다면  $M^2$ 은 점근적으로( $n_{11k} \rightarrow \infty$  이거나  $K \rightarrow \infty$ ) 자유도가 1인 카이제곱분포를 따른다는 사실이 알려져 있다.

CMH 검정의 장점은 가능도비통계량(likelihood ratio statistic)을 이용하는 검정과 달리 각 분할표의 칸의 도수가 작은 경우(sparse tables)에도 적용 가능하다는 점이다. 예를 들어 각 분할표(또는 층)가 대응표본에 의해 얻어진 경우, 즉  $n_{1+k} = n_{2+k} = 1$ 일 때에도 층의 수  $K$ 가 크면  $M^2$ 은 근사적으로 자유도 1인 카이제곱분포를 따른다.

CMH검정은 X와 Y의 상관관계가 Z의 수준에 관계없이 일관되게 양이거나 일관되게 음인 경우,  $n_{11k} - m_{11k}$ 의 부호가 한 쪽 방향으로 치우치게 되고, 따라서  $M^2$ 의 값이 크게 되어 조건부 독립이라는 영가설을 기각하게 된다. 하지만 X와 Y의 관련성이 층에 따라 크게 변하여  $n_{11k} - m_{11k}$ 의 부호가 심하게 교차하는 경우, 즉 X, Y, Z간에 역동적인 3차 상호작용이 있을 때에는 적절하지 않은 검정이다. 본 논문에서는 층의 크기가 작고 X, Y, Z의 3차 상호작용이 있을 때에도 적용할 수 있는 검정통계량을 제안하고, 새로운 통계량을 이용한 검정과 CMH검정의 검정력을 비교해 보고자 한다.

제2절에서는 CMH 통계량  $M^2$ 에 대한 대안 통계량을 제안하였고, 제3절에서 제1종 오류의 확률과 검정력으로 두 검정통계량을 비교해 보고, 실제 자료에 적용시켜 보았다. 그리고 제4절에서 비교의 결과를 요약하고 앞으로 연구되어야 할 과제를 지적하였다.

## 2. 조건부 독립성 검정을 위한 대안 통계량

Cochran-Mantel-Haenszel의 조건부 독립성 검정은 X, Y, Z의 3차 상호작용이 있을 때에는 적절하지 않다. 즉, X와 Y의 관련성이 층에 따라 크게 변하여

$$n_{11k} - m_{11k}, k = 1, \dots, K$$

의 부호가 교차하는 경우에, 조건부 독립이 아님에도 불구하고 CMH 검정통계량  $M^2$ 의 값이 작게 되어 조건부 독립이라는 영가설을 제대로 기각하지 못하는 문제점이 있다. 따라서  $n_{11k} - m_{11k}$ 의 부호에 관계없이 그 크기에만 의존하는 통계량  $M_A^2$ 을 대안 검정통계량으로 제안한다.

$$\begin{aligned} M_A^2 &= \frac{[\sum_{k=1}^K |n_{11k} - m_{11k}| - E(\sum_{k=1}^K |n_{11k} - m_{11k}|)]^2}{V(\sum_{k=1}^K |n_{11k} - m_{11k}|)} \\ &= \frac{[\sum |n_{11k} - m_{11k}| - \sum E|n_{11k} - m_{11k}|]^2}{\sum V|n_{11k} - m_{11k}|} \end{aligned} \quad (2.1)$$

위 식에서  $E(n_{11k})$ 와  $V(n_{11k})$ 은 각각 식 (1.1)과 (1.2)에서 정의하였다. 한편

$$\begin{aligned} V|n_{11k} - m_{11k}| &= E(n_{11k} - m_{11k})^2 - (E|n_{11k} - m_{11k}|)^2 \\ &= V(n_{11k}) - (E|n_{11k} - m_{11k}|)^2 \end{aligned}$$

이므로, 식 (2.1)에서 추가적으로 계산이 필요한 항은  $E|n_{11k} - m_{11k}|$ 뿐이다.

정리 2.1은 조건부 독립이라는 영가설 하에서 검정통계량  $M_A^2$ 의 근사적 분포가 자유도 1인 카이제곱분포임을 말해준다. 정리 2.1의 증명은 부록에 수록하였다.

정리 2.1  $K$ 개의  $2 \times 2$  분할표  $\{(n_{11k}, n_{12k}, n_{21k}, n_{22k}) : k = 1, \dots, K\}$ 에서  $\pi_{ij|k} = \pi_{i+|k}\pi_{+j|k}$ 라고 하자. 만약  $K \rightarrow \infty$ 일 때, 충분히 큰 상수  $B$ 에 대하여

$$n_k < B, \forall k$$

이고, 어떤 상수  $\epsilon$ 에 대하여

$$\#\{k : 0 < \epsilon \leq \pi_{11|k} \leq 1 - \epsilon < 1\} \rightarrow \infty$$

이면, 행과 열의 합  $\{(n_{1+k}, n_{+1k}, n_{2+k}, n_{+2k}) : k = 1, \dots, K\}$ 가 주어졌을 때

$$M_A^2 = \frac{\sum_{k=1}^K |n_{11k} - m_{11k}| - \sum_{k=1}^K E|n_{11k} - m_{11k}|}{(\sum_{k=1}^K V|n_{11k} - m_{11k}|)^{1/2}}$$

의 분포는  $K \rightarrow \infty$ 일 때 표준정규분포에 수렴한다. 여기서  $\#$ 는 집합의 원소의 개수를 나타낸다.

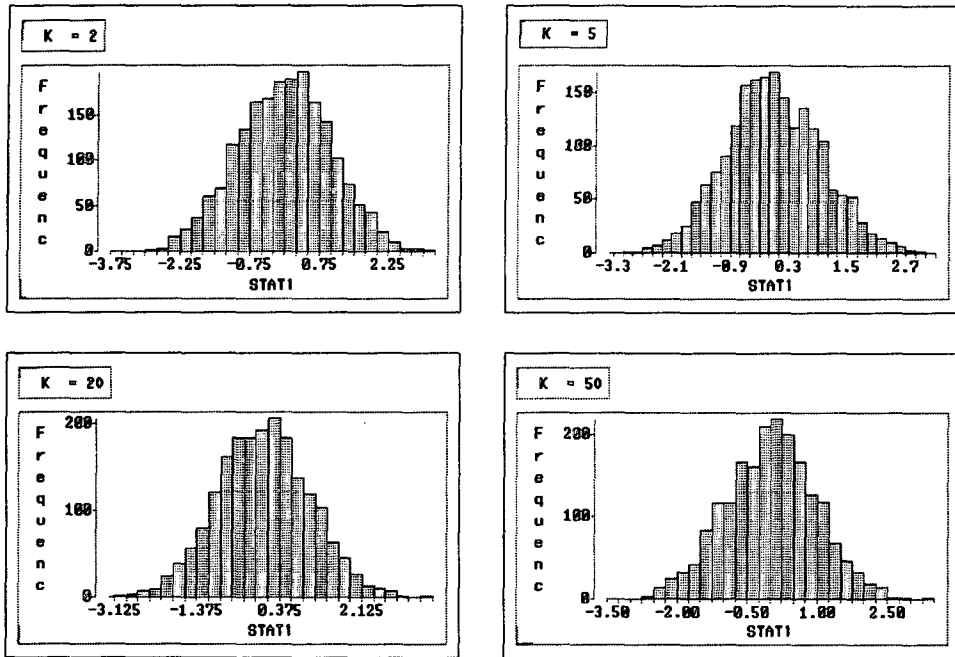


그림 2.1:  $M$ 의 분포를 나타내는 히스토그램

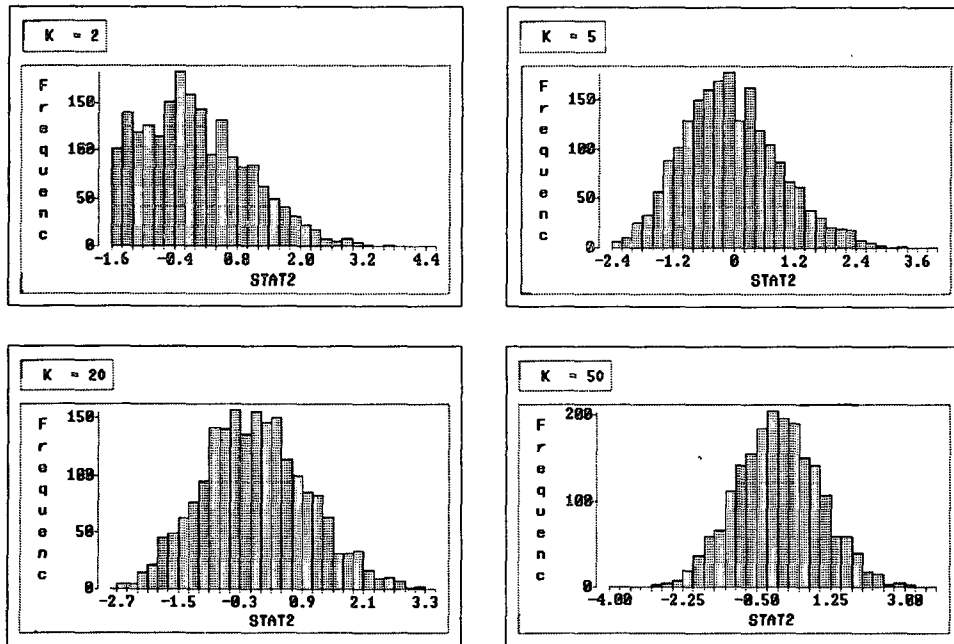


그림 2.2:  $M_A$ 의 분포를 나타내는 히스토그램

층의 수  $K$ 가 크지 않을 때 카이제곱분포가 얼마나 좋은 근사분포인지 알아보기 위해  $K=2, 5, 20, 50$ 에 대해 모의실험을 하였다. 먼저 각 층의 크기  $n_{++k} (k=1, \dots, K)$ 에 변화를 주기 위해 이항분포  $B(20, 0.5)$ 에서 난수를 생성하여  $n_{++k}$ 의 값으로 정하였다. 모든  $k$ 에 대하여  $\pi_{+1|k}, \pi_{1+|k}$ 를 0.5로 고정하고,  $\pi_{11|k} = \pi_{+1|k}\pi_{1+|k} = 0.25$ 로서 조건부 독립성을 만족하게 고정한다. 나머지 칸의 확률  $\pi_{12|k}, \pi_{21|k}, \pi_{22|k}$ 도 조건부 독립성을 만족하게 모두 0.25로 정한 다음, 다항확률분포  $\text{multinomial}(n_{++k}, \pi_k)$ ,  $\pi_k=(1/4, 1/4, 1/4, 1/4)$ 를 따르는 난수를 생성하여  $k$ 번째  $2 \times 2$ 분할표를 만든다. 이렇게 만든  $K$ 개의 분할표로부터  $M = [\sum n_{11k} - \sum m_{11k}] / [\sum V(n_{11k})]^{1/2}$ 과  $M_A$ 를 계산한다. 이와 같은 작업을 2,000번 반복하여 얻은  $M$ 과  $M_A$ 의 값들로서 그림 2.1, 그림 2.2와 같이 히스토그램을 그려보았다. 모의 실험을 위해 SAS를 사용하였으며,  $E|n_{11k} - m_{11k}|$ 의 값은 SAS의 함수 PROBHYPR을 이용하여 구하였다.

$M$ 과  $M_A$ 의 분포가 표준정규분포에 가까우면  $M^2$ 과  $M_A^2$ 의 분포가 카이제곱분포에 가까울 것이다. 히스토그램으로 나타낸 자료에 대해 정규성검정(SAS에서 Shapiro-Wilks 검정 이용)을 해 본 결과, 모든 층의 수  $K$ 에 대해  $M$ 의 분포는  $p$ 값이 0.15보다 커서 정규성에 별 문제가 없었다. 한편  $M_A$ 의 분포는 층의 수가 2, 5일 때는  $p$ 값이 0.01보다 작았고, 10에서 40 사이에는 0.01에서 0.1사이를, 그리고 층의 수가 50일 때는  $p$ 값이 0.15보다 컸다. 따라서 CMH 통계량  $M^2$ 은 문제가 없으나,  $M_A^2$ 은 층의 수가 작은 경우  $M_A$ 가 오른쪽 꼬리가 긴 분포를 가지므로,  $M_A^2$ 의 분포로서 자유도가 1인 카이제곱분포는 좋은 근사분포가 아님을 알 수 있다. 제 3절에서  $M_A^2$ 의 분포로 카이제곱분포를 가정했을 때 제 1종 오류의 확률에 관한 구체적 논의를 하였다.

### 3. 조건부 독립성 검정의 비교

#### 3.1. 제 1종 오류의 확률 비교

두 통계량  $M^2, M_A^2$ 을 이용한 조건부 독립성 검정의 제 1종 오류의 확률을 모의실험을 통해 비교하여 보자. 먼저,  $K$ 개의 층의 크기  $n_{++k} (k=1, \dots, K)$ 는 이항분포  $B(n, 0.5)$ 에서 생성한 난수로 정하였는데, 층의 평균 크기  $0.5n$ 의 값을 2, 3, 5, 10, 20, 30, 50, 70, 100, 200, 300, 500과 같이 변화시켜, 제 1종 오류의 확률이 층의 평균 크기에 따라 어떻게 변하는가를 보았다. 층의 수  $K$ 의 값도 다양하게 변화시켜 보았으나 여기서는  $K=5, 20, 50$ 일 때의 결과만 수록하였다. 분할표의 확률구조는 제 2절에서와 같이, 각 행과 열의 주변확률이 0.5이며 조건부 독립성을 만족하게 하였다. 각 검정이 범하게 되는 제 1종 오류 확률을 추정하기 위해, 정해진 각각의  $K$ 와 층의 평균 크기에 대해 2,000번 반복 실험 하였으며, 유의수준  $\alpha$ 는 0.05로 하였다. 그림 3.1에서  $M^2$ 을 이용한 CMH 검정의 제 1종 오류 확률은 ...로,  $M_A^2$ 의 경우는 \*\*\*로 표시하였다. 가로축은 층의 평균 크기로서 로그축적으로 나타내었다. 점선은  $0.05 \pm 3\sigma$ 를 나타내는데  $\sigma = \sqrt{0.05(1-0.05)/2000} = 0.00487$ 이다.

그림 3.1에서 층의 수와 층의 크기가 커질수록 두 검정의 제 1종 오류 확률에 큰 차이가 없어짐을 알 수 있다. 그리고 층의 수  $K$ 가 크면 층의 크기에 관계없이  $M_A^2$ 의 제 1종 오류 확률에는 문제가 없다.

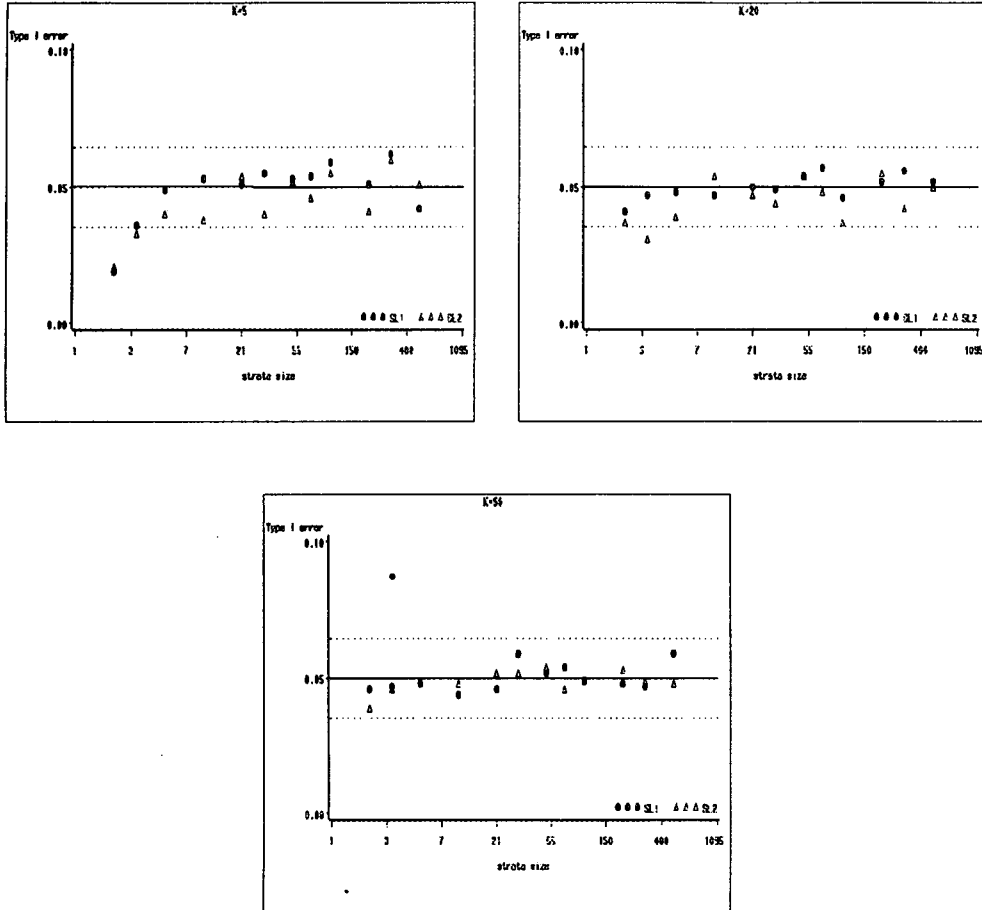


그림 3.1:  $M^2$ ,  $M_A^2$ 을 이용한 두 검정의 제 1 종 오류 확률

### 3.2. 검정력 비교

**< X와 Y의 상관관계가 일관된 방향일 때 >**

모의실험을 통해 X와 Y가 일관되게 양의 상관관계를 가질 때 검정력을 비교하여 보자. 0.25에서 0.5사이에서 균일한 분포(uniform distribution)를 갖는 난수를 생성하여  $\pi_{11|k}$  ( $k = 1, \dots, K$ )의 값으로 정하고,  $\pi_{12|k} = 0.5 - \pi_{11|k}$ ,  $\pi_{21|k} = 0.5 - \pi_{11|k}$ ,  $\pi_{22|k} = 1 - \pi_{11|k} - \pi_{12|k} - \pi_{21|k}$ 와 같이 정한다. 그러면 X와 Y가 일관되게 양의 상관관계를 가지는 확률구조가 된다. 이러한 확률구조를 갖는 다항분포에서 난수를 생성하여 분할표를 얻는다. 층의 평균크기와 층의 수를 변화시켜가며 각각에 대해 2000번씩 반복하여 검정력을 구해 보았다. 검정의

유의수준은 0.05로 하였으며 2000번 중에서 영가설을 기각하게 되는 확률을 검정력으로 정의하였다.

그림 3.2는 각각  $M^2$ 과  $M_A^2$ 을 이용한 검정의 검정력을 비교하기 위해 검정력의 비율

$$\text{ratio1} = \frac{M_A^2 \text{의 검정력}}{M^2 \text{의 검정력}}$$

을 층의 수  $K$ 가 5, 20, 50일 때, 그리고 층의 평균 크기의 여러 값에 대해 그린 것이다. 그림 3.2에서 보면,  $X$ 와  $Y$ 가 일관되게 양의 상관관계를 가질 때, 예상대로  $M^2$ 의 검정력이  $M_A^2$ 의 검정력보다 높으나, 층의 수와 층의 크기가 클수록 두 검정력의 차이는 없어짐을 알 수 있다. 층의 수가 5일 때는 층의 크기가 20이 넘어야 검정력의 비가 1에 수렴하지만 층의 수가 50일 때는 층의 크기가 3만 되어도 1에 수렴하는 것으로 보아 검정력이 층의 수에 영향을 많이 받음을 알 수 있다.

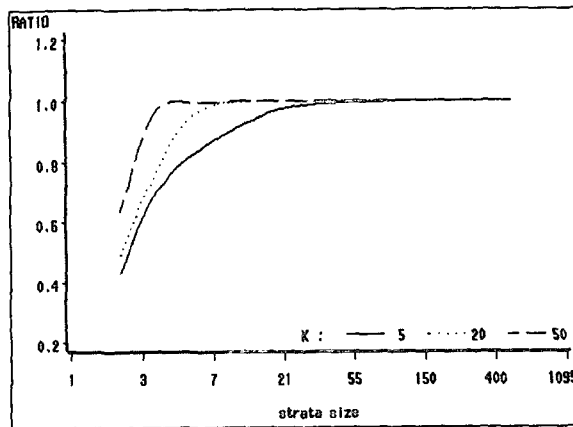


그림 3.2:  $X$ 와  $Y$ 가 일관되게 양의 상관관계를 가질 때 ( $M_A^2$ 의 검정력)/( $M^2$ 의 검정력)

**( X와 Y의 상관관계의 방향이 변할 때 )**

$X$ 와  $Y$ 의 상관관계가  $Z$ 의 수준에 따라 역동적으로 변하여,  $n_{11k} - m_{11k}$ 의 부호가 교차할 때, 과연  $M_A^2$ 을 이용한 검정이 기대한 것처럼 더 좋은 검정력을 갖는지 모의실험을 통해 알아보자.  $2 \times 2$  분할표에서 주변확률  $\pi_{1+|k}$ ,  $\pi_{2+|k}$ ,  $\pi_{+1|k}$ ,  $\pi_{+2|k}$ 를 0.5로 고정시키고  $\pi_{11|k}$ 를 0에서 0.5사이에서 같은 확률로 추출한다. 이러한 확률구조로부터 생성된  $K$ 개의 분할표의  $X$ - $Y$ 상관관계는 양과 음을 교차하게 된다.  $K$ 개의 분할표로부터 두 검정통계량의 값을 계산하고 유의수준 0.05에서 영가설 기각 여부를 결정한다. 이러한 작업을 2,000번 반복하여 두 검정통계량의 검정력을 비교하였다.

그림 3.3은 층의 수를 5, 20, 50로 하고, 층의 평균 크기를 2, 3, 5, 10, 20, 30, 50, 70, 100,

200, 300, 500로 했을 때의 검정력의 비

$$\text{ratio2} = \frac{M^2\text{의검정력}}{M_A^2\text{의검정력}}$$

을 나타낸 것이다. 100과 같이 큰 층의 크기에 대해서도 층의 수에 관계없이 검정력의 비가 0.8에 미치지 못하므로  $M_A^2$ 의 검정력이  $M^2$ 의 검정력보다 훨씬 좋다고 볼 수 있다. 따라서 X와 Y의 상관관계가 Z의 수준에 따라 역동적으로 변하여,  $n_{11k} - m_{11k}$ 의 부호가 교차할 때에는 검정통계량  $M_A^2$ 에 의한 검정이 적절함을 알 수 있다.

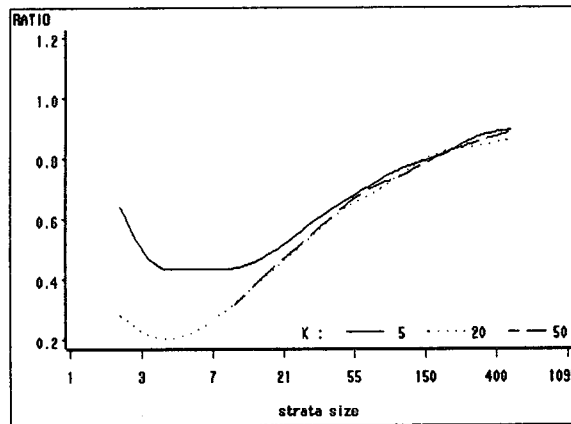


그림 3.3: X-Y 상관관계가 변할 때 ( $M^2$ 의 검정력)/( $M_A^2$ 의 검정력)

### 3.3. 실제 자료에서의 계산 예

Mantel(1963)의 연구에서, 토끼를 대상으로 베타용혈성 연쇄구균(*β-hemolytic streptococci*)을 주입(injection)한 후 페니실린을 즉시 주사한 경우와 1시간 30분 후에 주사한 경우의 반응을 비교하고자 하였다. 반응은 치료가 된 경우와 죽은 경우 두 가지로 나누고 페니실린의 양을 1/8, 1/4, 1/2, 1, 4로 변화시켜 실험하였다.

페니실린의 주사방법(X)과 반응(Y)의 주사량(Z)에 대한 조건부 독립성을  $M^2$ 으로 검정하여 보면,

$$M^2 = \frac{[\sum_1^5 (n_{11k} - m_{11k})]^2}{\sum_1^5 V(n_{11k})} = 5.66$$

이므로, 자유도 1인 카이제곱분포로부터 p값은 0.017로 계산된다. 그런데 만약 페니실린의 양에 따라 주사방법의 효과가 달라질 수 있다면, 다시 말해 페니실린의 주사량이 적을 때에는 즉시 주사하는 방법이 치료효과가 크지만, 주사량이 많을 때에는 역으로 1시간 반 후에 주사하는 방법이 더 치료효과가 크다면,  $M^2$ 을 이용한 CMH검정은 효과적인 검정이 되지 못한다. 다음 자료는 그러한 경향을 보여준다.



표 3.1: 페니실린 주사방법에 따른 주사효과 (출처: Agresti(1990)에서 Mantel(1963)의 자료를 인용한 것을 재인용)

페니실린 양	주사 시간	반응		$n_{11k} - m_{11k}$	$V(n_{11k})$	$ n_{11k} - m_{11k}  - E n_{11k} - m_{11k} $	$V n_{11k} - m_{11k} $
		치료	사망				
1/8	즉시	0	6	0	0	0	0
	1시간 반	0	5				
1/4	즉시	3	3	1.5	0.614	0.818	0.149
	1시간 반	0	6				
1/2	즉시	6	0	2	0.728	1.394	0.36
	1시간 반	2	4				
1	즉시	5	1	-0.5	0.25	0	0
	1시간 반	6	0				
4	즉시	2	0	0	0	0	0
	1시간 반	5	0				

이러한 경우에 적절한 검정통계량  $M_A^2$ 으로 조건부 독립성을 검정하여 보면,

$$M_A^2 = \frac{[\sum_1^5 (|n_{11k} - m_{11k}| - E|n_{11k} - m_{11k}|)]^2}{\sum V(|n_{11k} - m_{11k}|)} = 9.62$$

로서 (p값=0.002), 독립이 아니라는 증거가 더 크게 된다.

위 자료에서 조건부 독립을 기각할 수 있는 증거의 정도는 다르지만 유의수준 0.05에서 결론은 같다. 그런데, 만약 나머지 자료는 같고 페니실린의 양이 4일 때만 다음과 같이 관찰되었다면,

	치료	사망
즉시	1	1
1시간 30분	5	0

$M^2=2.91$ (p값=0.08)이고,  $M_A^2=11.61$ (p값=0.0007)로서 검정통계량에 따라 다른 결론을 얻게 된다. 따라서 X, Y, Z의 3차 상호작용에 대한 사전지식이 없을 때,  $M_A^2$ 을 이용한 검정이 보완적으로 실시되어야 함을 알 수 있다.

#### 4. 결론

$M^2$ 을 이용한 CMH검정은, X, Y, Z의 3차 상호작용이 없는 경우, 층의 수와 층의 크기에 관계없이 적절한 검정이다. 하지만 X와 Y의 상관관계가 Z의 수준에 따라 역동적으로 변할 때에는 CMH검정이 낮은 검정력을 갖는다는 것을 알 수 있었다. 이 때 새로이 제안한 검정통계량  $M_A^2$ 을 이용하면 검정력을 높일 수 있다. 그러나 층의 수가 작을 때, 자유도가 1인 카

이제 공급분포는  $M_A^2$ 의 영분포(null distribution)로서 좋은 근사가 되지 못함을 모의실험을 통해 확인할 수 있었다. 하지만 층의 수가 작을 때에도  $M_A^2$  검정의 제 1종 오류 확률이 정해진 유의수준을 넘지 않으며, X와 Y의 상관관계가 Z의 수준에 따라 역동적으로 변할 때에 검정력이 뛰어나므로,  $M_A^2$ 검정은 분명한 장점을 가진다.

표본크기가 크지 않은 K개의  $2 \times 2$ 분할표에 대해 조건부 독립성 검정을 할 때, X-Y상관관계가 Z의 수준에 관계없이 한 방향이라는 확신이 있다면 즉, 일관되게 양의 상관관계를 갖거나 일관되게 음의 상관관계를 갖는다면 CMH검정이 선호되어야 한다. 하지만 X-Y상관관계가 Z의 수준에 따라 그 방향까지 변할 수 있을 때나 3차 상호작용에 대한 사전지식이 없을 때  $M_A^2$ 을 이용한 검정을 먼저 실시해 보아야 할 것이다. 그래야만 역동적으로 존재하는 X-Y상관관계를 탐지할 수 있다.

한편, 표본크기가 클 때에는 로그선형모형을 이용하면 조건부 독립성을 포함하여 X, Y, Z에 대해 보다 많은 정보를 얻을 수 있을 것이다. 표본크기가 작을 때  $M_A^2$ 을 이용한 검정이 보완적으로 선행되어야 하는 것은, 마치 표본크기가 클 때 로그선형모형에서 통계량  $G^2[(XZ, YZ)|(XY, XZ, YZ)]$ 를 이용하여 X-Y의 조건부 독립성을 검정하기 전에  $G^2[(XY, XZ, YZ)|(XYZ)]$ 을 이용하여 X, Y, Z의 3차 상호작용의 유무를 먼저 판단해야 하는 것과 같다. (통계량  $G^2[(XZ, YZ)|(XY, XZ, YZ)]$ 와  $G^2[(XY, XZ, YZ)|(XYZ)]$ 의 정의는 Agresti(1990) 7.1절 참조)

끝으로 앞으로의 연구과제를 제시해 본다. 첫 번째 연구과제로서, 층의 수가 작을 때,  $M_A^2$ 통계량의 영분포로서 적절한 분포를 찾아내는 일이다. 이를 위해서는 영가설 하에서  $n_{11k}$  ( $k = 1, \dots, K$ )가 갖는 정확한 분포의 계산 알고리즘을 제시한 Mehta 등(1985, 1988)의 연구가 도움이 될 수 있을 것이다. 표본재추출(resampling)을 이용한 방법도 시도해 볼 예정이다. 두 번째 연구과제는 중도절단된 자료(censored data)에의 적용이다. 중도절단된 자료에서 두 모집단의 생존함수를 비교하는 문제에 통계량  $M^2$ 이 응용되는데, 여기서도 마찬가지로 통계량  $M_A^2$ 이 적용될 수 있다. 이 때 층의 수는 중도절단되지 않는 관측값의 개수와 같으므로  $M_A^2$ 의 영분포로서 카이제곱분포를 별 문제없이 쓸 수 있으리라 짐작된다. 시간에 따라 두 모집단의 고장률의 비가 상수가 아니거나 단조변화하지 않을 때,  $M_A^2$ 을 이용한 검정의 검정력이 더 높으리라 예상된다.

### 부록: 정리 2.1의 증명

증명:  $n_{++k} < B$ ,  $\forall k$ 이므로 각  $j = 1, \dots, K$ 에 대하여 다음을 만족하는 상수  $M_{Kj}$ 가 존재한다.

$$X_{Kj} = \frac{|n_{11j} - m_{11j}|}{[\sum_{k=1}^K V|n_{11k} - m_{11k}|]^{1/2}} \leq M_{Kj}$$

또한  $n_{++k} < B$ ,  $\forall k$ 이고  $\#\{k : 0 < \epsilon \leq \pi_{11|k} \leq 1 - \epsilon < 1\} \rightarrow \infty$ 이므로

$$\max_{1 \leq j \leq K} M_{Kj} \rightarrow 0$$

이다. 따라서 Chung (1974)의 201쪽 보조정리에 의하여  $M_A = \sum_{j=1}^K X_{Kj} - E(\sum_{j=1}^K X_{Kj})$ 는  $K \rightarrow \infty$ 일 때 표준정규분포에 수렴한다.  $\square$

## 참고문헌

- [1] Agresti (1990), *Categorical Data Analysis*, John Wiley & Sons: New York.
- [2] Chung, K. L. (1974), *A Course in Probability Theory*, 2nd ed. Academic Press: New York.
- [3] Cochran, W. G. (1954), Some methods of strengthening the common  $\chi^2$  tests. *Biometrics*. Vol. 10. 417-451.
- [4] Mantel, N. (1963), Chi-square tests with one degree of freedom: extension of the Mantel-Haenszel procedure. *Journal of American Statistical Association* . Vol. 58. 690-700.
- [5] Mantel, N., and Haenszel, W. (1959), Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of National Cancer Institute* . Vol. 22. 719-748.
- [6] Mehta, C. R., Patel, N. R., and Gray, R. (1985), Computing an exact confidence interval for the common odds ratio in several 2 by 2 contingency tables. *Journal of American Statistical Association* . Vol. 80. 969-973.
- [7] Mehta, C. R., Patel, N. R., and Senchaudhuri, P. (1988), Importance Sampling for Estimating Exact Probabilities in Permutational Inference. *Journal of American Statistical Association* Vol. 83. 999-1005.

[ 1997년 4월 접수, 1998년 2월 최종수정 ]

## A Study on Mantel-Haenszel Test of Conditional Independence

Ji-Hyun Kim <sup>1)</sup> Hyun Sun Lim <sup>2)</sup>

### ABSTRACT

Many epidemiological studies investigate whether an association exists between a binary risk factor  $X$  and a binary response variable  $Y$ . They analyze whether an observed association between  $X$  and  $Y$  persists when the level of another factor  $Z$  that might influence the association is controlled. This involves testing conditional independence of  $X$  and  $Y$  controlling for  $Z$ . The Mantel-Haenszel test is most widely used to test conditional independence for sparse tables. But if the association between  $X$  and  $Y$  varies along the levels of  $Z$ , Mantel-Haenszel test has a low power problem. In this study, we propose an alternative test procedure which overcomes the low power problem in that case. We find out the null distribution of the alternative test statistic and compare its performance with the Mantel-Haenszel test by simulation.

---

1) Associate Professor, Dept of Statistics, Soong Sil University, Sangdo 5-Dong, Dongjak-Ku, Seoul 156-743.

2) Assistant Instructor, Dept of Applied Statistics, Korea National Open University, Dongsoong-Dong, Jongro-Ku, Seoul 110-791.