

Interval Estimation of the Difference of two Population Proportions using Pooled Estimator

Chong Sun Hong¹⁾

ABSTRACT

In order to examine whether the difference between two point estimates of population proportions is statistically significant, data analysts use two techniques. The first is to explore the overlap between two associated confidence intervals. Second method is to test the significance which is introduced at most statistical textbooks under the common assumptions of consistency, asymptotic normality, and asymptotic independence of the estimates. Under the null hypothesis which is two population proportions are equal, the pooled estimator of population proportion is preferred as a point estimator since two independent random samples are considered to be collected from one population. Hence as an alternative method, we could obtain another confidence interval of the difference of the population proportions with using the pooled estimate. We conclude that, among three methods, the overlapped method is under-estimated, and the difference of the population proportions method is over-estimated on the basis of the proposed method.

Keywords : Confidence Intervals; Efficiency; Overlap; Pooled Estimator; Power; Test of Significance; Type 1 error.

1. 서론

두 모비율 p_1 과 p_2 사이의 차이에 대한 통계적 추론 방법 중에서 구간추정 방법들은 여러 방법이 있다. 자료분석에서 가장 많이 사용되는 방법으로는 각각의 모비율에 대한 95% 신뢰구간을 다음과 같이 구하여 그 신뢰구간이 중복(overlapping) 되는가를 검사하여 살펴보는 방법이 있다.

$$\hat{p}_1 \pm 1.96 \sqrt{\hat{p}_1(1-\hat{p}_1)/n_1}, \quad \hat{p}_2 \pm 1.96 \sqrt{\hat{p}_2(1-\hat{p}_2)/n_2}, \quad (1)$$

여기서 \hat{p}_1 과 \hat{p}_2 는 각각 p_1 과 p_2 의 추정량이며 n_1 과 n_2 는 각 확률표본의 크기이다. 만약 그 신뢰구간들이 중복되지 않는다면 모비율의 차이가 통계적으로 유의하다고 판단 내릴 수 있으며, 중복된다면 차이가 유의하지 않다고 판단할 수 있다. 이 방법은 각 구간이 중복되는지의 여부를 검

1) Professor, Department of Statistics, Economics School, Sungkyunkwan University, Seoul, KOREA 110-745. E-mail : cshong@skku.ac.kr

사하면서 구간의 경계값(상한과 하한값)을 비교하는 것이 쉽기 때문에 사용하기가 편리한 방법이다. 그러나 이 방법은 잘못된 결론을 유도할 수 있기 때문에 사용을 자제하여야 한다는 연구 논문이 많다. 이와 관련된 최근의 논문들로는 Abergel와 그외(1998), Appleby와 Bell(2000), Chalmers와 Salmon(2000), Cole과 Blair(1999), Collins와 그외(2001), Debboun와 그외(2000), Djordjevic와 그외(2000), Eising과 Ohe(1998) Ezeamuzie와 그외(1999), Gotzsche(2000), Herr와 그외(1998), Inzelberg와 그외(2000), Kann와 그외(2000), Mancuso와 그외(2001), Martinez와 그외(1999), Mcbride(2000), Mehta와 그외(2000), Millikan와 그외(2000), Rahlfs(1997), Slavotinek와 그외(1999), Sont와 그외(2001), Tersmette와 그외(2001), Vatassery와 그외(1998), Wong와 그외(1999) 등과 같은 의학과 보건에 관한 학술지에 많이 게재되어 있다. 각각의 모비율의 95% 신뢰구간을 (1)식과 같이 구하여 구간이 중복되는지를 검사하는 방법(이하 중복방법)은 다음의 구간이

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96 \left(\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1} + \sqrt{\hat{p}_2(1-\hat{p}_2)/n_2} \right) \quad (2)$$

0값을 포함하는지를 살펴보는 것과 동일하다.

통계학 교재에 반드시 소개되는 가장 일반적인 방법으로는 두 모비율 차이 $p_1 - p_2$ 의 95% 신뢰구간을 다음과 같이 구하고 이 신뢰구간에 0값이 포함되는지의 여부를 관찰하는 방법(이하 비율차방법)이 있다.

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96 \sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2} . \quad (3)$$

위 신뢰구간의 상한이 양의 값이고 하한값이 음수이면, 구간에 0을 포함하고 이는 두 모비율의 차이가 통계적으로 유의하지 않다고 할 수 있다. 그리고 하한값이 양수인 경우에는 모비율 p_1 이 p_2 보다 크며, 상한값이 음수인 경우에는 모비율 p_1 이 p_2 보다 작다고 통계적인 의사결정을 내릴 수 있다.

Cole과 Blair(1999) 그리고 Schenker와 Gentleman(2001)은 일반적인 방법인 비율차의 신뢰구간을 살펴보는 방법인 비율차방법과 각각의 신뢰구간을 구하여 중복을 검사하는 방법인 중복방법을 비교하여 중복을 검사하는 방법이 일반적인 비율차방법보다 귀무가설 $H_0 : p_1 = p_2$ 을 기각하는 경우가 적기 때문에 보수적(conservative)이라고 하였다. 즉, 중복방법에 의하여 귀무가설을 기각시키는 경우에는 비율차방법에 의하여도 기각시킬 수 있으나 반대로 비율차방법에 의하여 귀무가설을 기각시키는 경우에는 중복방법에 의하여 기각시킬 수 없는 경우가 발생한다. 귀무가설이 사실일 때, 중복방법이 비율차방법보다 더욱 빈번하게 귀무가설을 기각하는 실수를 범한다고 하였다. 특히 Schenker와 Gentleman(2001)은 두 방법의 효율과 제1종의 오류의 확률 그리고 검정력 함수를 구하여 근사이론을 보다 발전시켰다.

두 모비율이 동일하다는 귀무가설이 사실인 경우에는 모비율의 분산 $\hat{p}_1(1-\hat{p}_1)/n_1$ 와 $\hat{p}_2(1-\hat{p}_2)/n_2$ 의 분자가 동일하게 된다. 따라서 귀무가설이 참인 경우에는 각각의 독립적인 표본에서 모비율 추정량의 기대값과 분산이 동일하기 때문에 두 개의 서로 다른 모집단에서 추출된

표본이 아니라 하나의 모집단에서 추출된 확률표본으로 간주하여야 한다. 그러므로 귀무가설 $H_0 : p_1 = p_2$ 이 참인 경우에 모비율의 합동추정량(pooled estimator)은

$$\hat{p}_0 = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} \quad (4)$$

이 되며, 모비율 차이의 합동분산 추정량은 다음과 같이 정의된다.

$$\hat{p}_0(1 - \hat{p}_0) \left(\frac{1}{n_1} + \frac{1}{n_2} \right). \quad (5)$$

합동추정량을 사용한 두 모비율 차이에 대한 95% 신뢰구간은 다음과 같다.

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96 \sqrt{\hat{p}_0(1 - \hat{p}_0) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}. \quad (6)$$

본 연구에서는 상기 언급한 두 가지 방법인 중복방법과 비율차방법과 새로 제시한 합동추정량을 사용한 구간추정 방법(이하 합동방법)을 비교 분석하고자 한다. 2절에서는 연구에 필요한 개념을 정의하고 가정을 설정한다. 그리고 3절에서는 세 가지 방법에 의한 신뢰구간의 효율성을 비교한다. 4절에서는 세 가지 추정방법들의 검정력 함수를 구하고 이 함수들을 비교 분석하여 귀무가설이 참일 때 제1종의 오류와 전반적인 검정력을 토론한다. 마지막 5절에서는 세 방법의 평가를 논의한다.

2. 정의와 가정

본 논문에서 연구하는 추정량의 일치성과 근사적 정규성 그리고 독립성을 가정한다. 이를 상세히 하기 위하여 두 모집단의 비율 p_1 과 p_2 의 차이에 관한 추론에 관하여 다음을 정의하고 가정하자. \hat{p}_1 과 \hat{p}_2 는 각각 p_1 과 p_2 의 일치추정량(consistent estimator)이며, \hat{p}_1 과 \hat{p}_2 의 표준오차(standard error) $\sqrt{\hat{p}_1(1 - \hat{p}_1)/n_1}$ 과 $\sqrt{\hat{p}_2(1 - \hat{p}_2)/n_2}$ 는 대응하는 각각의 모비율의 표준오차 $\sqrt{p_1(1 - p_1)/n_1}$ 과 $\sqrt{p_2(1 - p_2)/n_2}$ 의 일치추정량이라고 하자. 그리고 일치추정량 \hat{p}_1 과 \hat{p}_2 는 근사적으로 정규분포를 따르며 각각 독립을 가정한다. 즉 표본크기가 커질수록 다음과 같이 수렴한다.

$$\begin{aligned} \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{p_1(1 - p_1)}} &\rightarrow 1 \quad \text{in probability} \\ \sqrt{\frac{\hat{p}_2(1 - \hat{p}_2)}{p_2(1 - p_2)}} &\rightarrow 1 \quad \text{in probability} \end{aligned}$$

$$\frac{\hat{p}_1 - p_1}{\sqrt{p_1(1-p_1)/n_1}} \rightarrow Z_1 \quad \text{in distribution}$$

$$\frac{\hat{p}_2 - p_2}{\sqrt{p_2(1-p_2)/n_2}} \rightarrow Z_2 \quad \text{in distribution}$$

여기서 Z_1 과 Z_2 는 독립적인 표준정규분포를 따르는 확률변수이다.

위와 같은 가정을 이용하여 간단한 극한이론(limiting theory)으로부터 다음의 세 통계량은 극한 표준정규분포를 따른다는 것을 유도할 수 있다.(Cramer 1946, pp. 254-255 참조.)

$$\frac{\hat{p}_1 - p_1}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1}}, \quad \frac{\hat{p}_2 - p_2}{\sqrt{\hat{p}_2(1-\hat{p}_2)/n_2}}, \quad \text{그리고} \quad \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}}.$$

또한 귀무가설 $H_0 : p_1 = p_2$ 하에서 다음과 같이 정의된 통계량도 극한 정규분포를 따른다는 것을 유도할 수 있다. (홍종선과 최현집 1999, pp. 19-23 참조. 특히 2×2 분할표의 독립성 검정통계량과 동질성 검정통계량인 Pearson X^2 통계량이 Z_0 의 제곱과 동일하다는 것을 보였다.)

$$Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}. \quad (7)$$

3. 신뢰구간 효율성 비교

유의수준 5%로 하여 두 모비율의 차이에 대한 95% 신뢰구간을 계산하는 방법들은 (2)식과 (3)식 그리고 (6)식에 의한 세 종류로 나열할 수 있다. (2)식에 의한 방법은 각 모비율의 신뢰구간이 중복되는지를 검정하는 방법이기 때문에 중복방법(overlapped method)라고 하고, (3)식에 의한 방법은 일반적인 방법으로 모비율 차이에 대한 신뢰구간에 의한 방법이므로 비율차방법(proportion difference method)라고 한다. 그리고 (6)식에 의한 방법은 귀무가설이 사실일 때 모비율의 합동추정량과 합동분산 추정량을 사용한 모비율 차이의 신뢰구간 방법이므로 합동방법(pooled method)이라고 하자.

(2), (3), 그리고 (6)식을 통해서 살펴보면, 이들 세 가지 방법의 차이는 다음과 같은 각각의 신뢰구간 폭에 기인하기 때문에

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}},$$

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}},$$

$$\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

의 비교로 국한된다.

중복방법과 비율차방법 사이의 비교연구를 수행한 Schenker와 Gentleman(2001)은 중복방법의 표준오차가 비율차방법의 표준오차보다 크거나 같다고 상세히 논의하였다. 그리고 합동방법과 비율차방법의 표준오차 추정량들을 비교하여 보자. 비교를 간단히 하기 위하여 두 표본의 크기를 동일하다고 가정하고 (즉, $n_1 = n_2 = n$) 표준오차의 제곱인 표본분산 추정량을 고려하여 그 차이를 구하면 다음과 같다.

$$\begin{aligned} & \widehat{p}_0(1-\widehat{p}_0)\left(\frac{2}{n}\right) - \frac{\widehat{p}_1(1-\widehat{p}_1) + \widehat{p}_2(1-\widehat{p}_2)}{n} \\ &= \left(\frac{\widehat{p}_1 + \widehat{p}_2}{n}\right)\left(1 - \frac{\widehat{p}_1 + \widehat{p}_2}{2}\right) - \frac{\widehat{p}_1(1-\widehat{p}_1) + \widehat{p}_2(1-\widehat{p}_2)}{n} \\ &= \frac{2}{n}(\widehat{p}_1 - \widehat{p}_2)^2 \geq 0 \end{aligned}$$

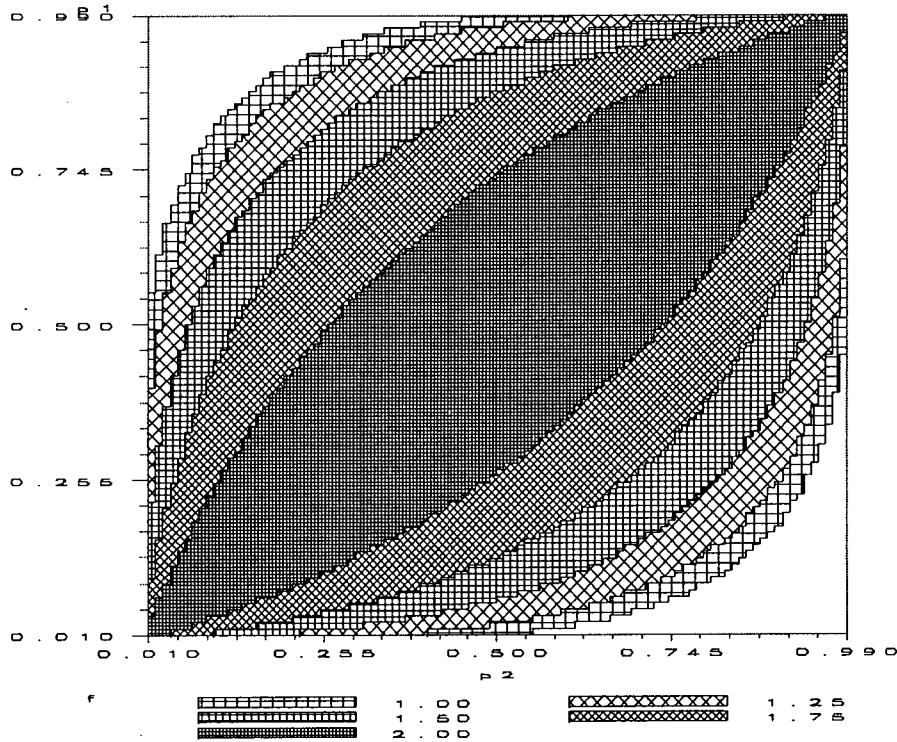
따라서 합동방법의 표준오차 추정량이 비율차방법의 표준오차 추정량보다 크거나 같다. 또한 중복방법의 표준오차 추정량과 합동방법의 표준오차 추정량과의 비율은 다음과 같다.

$$\left(\sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n}} + \sqrt{\frac{\widehat{p}_2(1-\widehat{p}_2)}{n}} \right) / \sqrt{\left(\frac{\widehat{p}_1 + \widehat{p}_2}{n}\right)\left(1 - \frac{\widehat{p}_1 + \widehat{p}_2}{2}\right)} \quad (8)$$

이 비율이 1보다 크거나 같을 때의 \widehat{p}_1 과 \widehat{p}_2 는 또한 <그림 1>과 같은 경우이다. <그림 1>은 (8)식의 비율의 크기를 나타낸 그림으로, 그 비율이 1보다 크거나 같을 때에 대응하는 모비율의 추정량 \widehat{p}_1 과 \widehat{p}_2 의 영역을 나타내었다. 그리고 이 비율이 1보다 작은 경우에는 왼쪽 상단과 오른쪽 하단 부분과 같이 여백으로 표현하였다. <그림 1>을 살펴보면, \widehat{p}_1 과 \widehat{p}_2 의 값이 <그림 1>에서 나타난 범위 안에 있을 때 합동방법의 표준오차 추정량은 중복방법의 표준오차 추정량보다 작거나 같다. 특히 \widehat{p}_1 과 \widehat{p}_2 의 값이 유사할 때 (즉, $\widehat{p}_2 = \widehat{p}_1$ 의 직선 근처의 값) 대응하는 표준오차 추정량들의 차이가 크게 발생하여 중복방법이 표준오차 추정량은 합동방법의 표준오차 추정량의 $\sqrt{2}$ 배가 되며, \widehat{p}_1 과 \widehat{p}_2 의 차이가 커질수록 표준오차 추정량의 차이가 작아짐을 알 수 있다. 그러나 \widehat{p}_1 과 \widehat{p}_2 값의 차이가 매우 큰 경우에는 중복방법의 표준오차 추정량의 값보다 합동방법의 표준오차 추정량이 커지는 반대 현상이 나타난다. 그런데 모비율 추정량 \widehat{p}_1 과 \widehat{p}_2 의 차이가 큰 경우에는 귀무가설 $H_0 : p_1 = p_2$ 을 검정하여 귀무가설을 기각시키는 결과를 유도하는 과정이 무의미하다. 즉, 이런 경우의 검정을 수행하는데 통계적 의미를 부여할 필요가 없다. 따라서 <그림 1>과 같이 \widehat{p}_1 과 \widehat{p}_2 의 값의 차이가 큰 경우를 제외한 대부분의 경우에서와 같은 경우에는 다음과 같은 부등식이 성립한다.

$$\text{중복방법의 표준오차} \geq \text{합동방법의 표준오차} \geq \text{비율차방법의 표준오차} \quad (9)$$

위 (9)식과 같은 부등식에서 \hat{p}_1 과 \hat{p}_2 이 유사한 값을 가질 때는 중복방법의 표준오차는 합동방법의 표준오차의 $\sqrt{2}$ 배 정도가 되며, 합동방법의 표준오차는 비율차방법의 표준오차와는 유사한 크기를 갖는다.



<그림 1> 중복방법과 합동방법의 표본분산 비율

귀무가설 $H_0 : p_1 = p_2$ 가 사실인 경우에는 독립적인 두 개의 확률표본에서 모비율 추정량의 기대값과 분산이 동일하다. 따라서 두 개의 확률표본이 서로 다른 모집단에서 추출한 것이 아니고 동일한 하나의 모집단에서 추출한 표본으로 간주하여야 한다. 그러므로 (4)식과 같은 합동추정량을 사용한 두 비율차의 구간추정방법을 사용하는 것이 바람직하다. 구간추정 방법 중에서 합동방법을 기준으로 한다면, 비율차방법에 의한 일반적인 구간추정방법은 과소추정(under estimated) 되었고, 중복방법에 의한 구간추정방법은 과대추정(over estimated) 되었다고 볼 수 있다.

4. 검정력 함수 비교

2절에서 논의했듯이 합동추정량을 사용한 두 비율차이의 구간추정방법이 제일 바람직하므로 귀무가설 $H_0 : p_1 = p_2$ 하에서의 검정통계량 $Z_0 = (\hat{p}_1 - \hat{p}_2) / \sqrt{2\hat{p}_0(1-\hat{p}_0)/n}$ 을 “표준화된 차이”라고 하자. 설정된 표준화된 차이를 이용하여 다양한 차이의 값에 따라 세 가지 구간추정방법의

검정력 함수의 비교가 가능하게 한다.(Cox와 Hinkley, 1974, pp. 317-318 참조.) 유의수준 $\alpha = 0.05$ 에서 검정을 수행할 때, 표준화된 차이값 d 에 대하여 세 가지 추정방법의 근사 검정력 함수는 다음과 같다.

1) 합동방법(pooled method)의 근사 검정력 함수

$$P[|Z_0 - d| \geq 1.96] = \Phi(-1.96 + d) + \Phi(-1.96 - d) \quad (10)$$

2) 비율차방법(proportional difference method)의 근사 검정력 함수

$$\begin{aligned} & \Phi\left(-1.96 \frac{\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}}{\sqrt{2\hat{p}_0(1-\hat{p}_0)}} + d\right) \\ & + \Phi\left(-1.96 \frac{\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}}{\sqrt{2\hat{p}_0(1-\hat{p}_0)}} - d\right) \end{aligned} \quad (11)$$

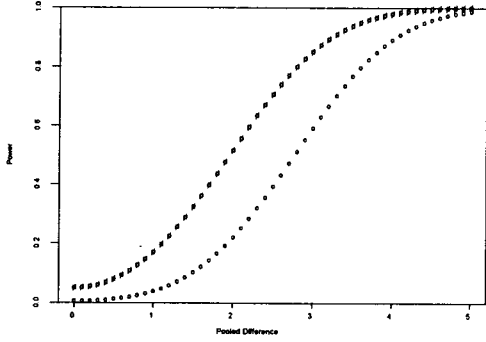
3) 중복방법(overlapped method)의 근사 검정력 함수

$$\begin{aligned} & \Phi\left(-1.96 \frac{\sqrt{\hat{p}_1(1-\hat{p}_1)} + \sqrt{\hat{p}_2(1-\hat{p}_2)}}{\sqrt{2\hat{p}_0(1-\hat{p}_0)}} + d\right) \\ & + \Phi\left(-1.96 \frac{\sqrt{\hat{p}_1(1-\hat{p}_1)} + \sqrt{\hat{p}_2(1-\hat{p}_2)}}{\sqrt{2\hat{p}_0(1-\hat{p}_0)}} - d\right) \end{aligned} \quad (12)$$

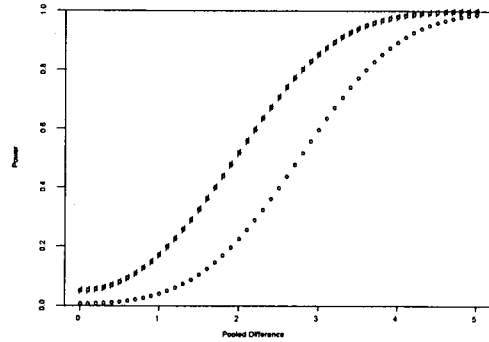
여기서 $\Phi(\cdot)$ 는 표준정규누적분포함수이다.

제1종 오류의 확률은 표준화된 차이가 $d=0$ 값을 가질 때 (즉, 귀무가설이 사실일 때) 근사 검정력 함수의 값이다. 따라서 위 식 (10), (11) 그리고 (12)에서 d 의 값을 0으로 설정하여 제1종 오류의 확률을 계산할 수 있으며, <그림 2>에서 표준화된 차이가 0일 때 그 값들을 비교할 수 있다. 합동방법에서의 제1종 오류는 항상 유의수준인 $\alpha=0.05$ 의 값을 가진다. 반면에 비율차방법의 제1종 오류는 0.05보다 크거나 같으며, 중복방법의 제1종 오류의 확률은 대부분의 경우에 0.05보다 작다. 그리고 두 모비율의 차이가 매우 큰 경우(예를 들어 $p_1=0.1$, $p_2=0.9$)에는 0.05보다 큰 값을 가질 수도 있다.(물론 이 경우에 검정을 수행하는데 의미는 없다.)

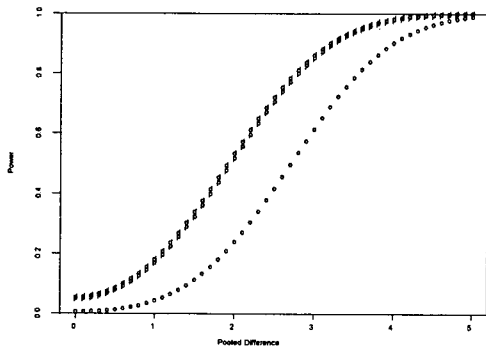
<그림 2>는 p_1 과 p_2 의 여러 값에 따라 표준화된 거리의 함수로서 세 가지 방법의 근사 검정력 함수를 나타내었다. p_1 과 p_2 의 값이 유사할 때는 비율차방법과 합동방법이 유사하게 높은 검정력을 갖고 있으며 중복방법의 검정력이 낮다. 그리고 p_1 과 p_2 의 값 차이가 클수록 비율차방법의 검정력이 제일 높고 합동방법의 검정력이 세 방법 중에서 가운데 위치하고 중복방법의 검정력이 낮음을 파악할 수 있다. 그리고 p_1 과 p_2 의 값의 변화에 따라 표준오차 추정량의 변동만큼 <그림 2>에서의 검정력 함수도 변화함을 알 수 있다.



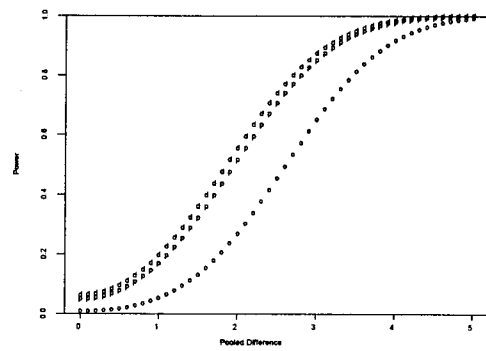
(1) $p_1 = 0.5, p_2 = 0.5$



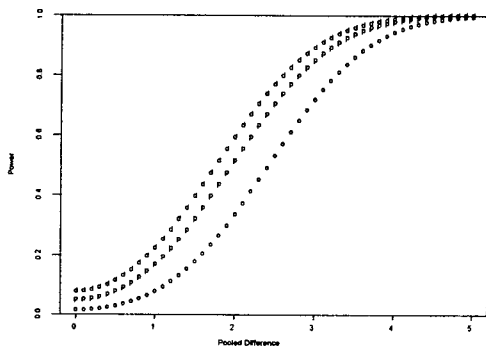
(2) $p_1 = 0.5, p_2 = 0.6$



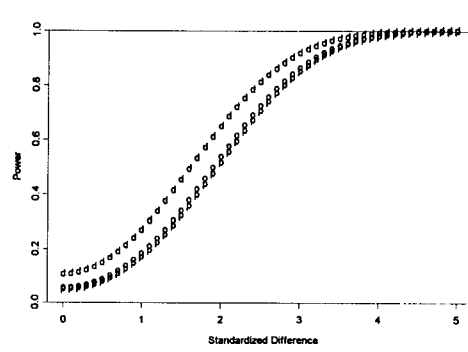
(3) $p_1 = 0.5, p_2 = 0.7$



(4) $p_1 = 0.5, p_2 = 0.8$



(5) $p_1 = 0.5, p_2 = 0.9$



(6) $p_1 = 0.5, p_2 = 1.0$

<그림 2> 근사 검정력 함수 (d= 비율차방법, o=중복방법, p=합동방법)

5. 결론

각 모비율의 신뢰구간을 구한 후 구간의 중복을 검사하는 방법은 간단하고 신뢰구간의 표나 그래프가 표현되었을 때 편리하게 이용할 수 있다. 이 방법은 빠르고 비교적 개략적인 탐색적 자료분석으로써 유용하다. 그러나 많은 통계학 교재에서 소개된 보편적인 방법인 비율차방법에 비교해서 중복방법은 표준오차가 과소추정됨으로 보수적이고 낮은 검정력 때문에 유의검정에 대해 바람직하다고 판단하기 어렵다. 그리고 일반적으로 많이 쓰이는 비율차방법은 귀무가설이 참이라는 조건하에서는 표준오차가 과소추정됨으로 신뢰구간의 폭이 좁아지며 검정력도 높다. 특히 두 비율 추정량 값의 차이가 증가할수록 과소추정되는 경향이 있으므로 공식적인 유의성 검정을 수행하기 위해 비율차방법을 사용할 때는 유의해야 한다. 그러므로 귀무가설 $H_0 : p_1 = p_2$ 하에서는 각 모비율의 분산의 분자가 동일한 형태를 갖는다. 따라서 귀무가설이 참인 경우에는 각각의 독립적인 표본에서 모비율 추정량의 기대값과 분산이 동일하기 때문에 두 개의 서로 다른 모집단에서 추출된 표본이 아니라 하나의 모집단에서 추출된 확률표본으로 간주해야 하기 때문에 합동추정량을 사용한 구간추정방법이 제일 적절하다고 할 수 있다.

참고문헌

- [1] 홍종선, 최현집 (1999). 「로그선형모형을 이용한 범주형자료분석」, 탐진.
- [2] Abergel, E., Linhart, A., Chatellier, G., Gariépy, J., Ducardonnet, A., Diebold. B., and Menard, J. (1998). Vascular and Cardiac Remodeling in World Class Professional Cyclists, *American Heart Journal*, Vol. 136, 818-823.
- [3] Appleby, J., and Bell, A. (2000). Reporting NHS Performance: How Did the Media Perform? *British Medical Journal*, Vol. 321, 248.
- [4] Chalmers, R. M., and Salmon, R. L. (2000). Primary Care Surveillance for Acute Bloody Diarrhea, Wales, *Emerging Infectious Diseases*, Vol. 6, 412-414.
- [5] Cole, S. R., and Blair, R. C. (1999). Overlapping Confidence Intervals, *Journal of the American Academy of Dermatology*, Vol. 41, 1051-1052.
- [6] Collins, S. L., Edwards, J. E., Moore, R. A., and McQuay, H. J. (2001). Single Dose Dextropropoxyphene, Alone and with Paracetamol (Acetaminophen), for Postoperative Pain (Cochrane Review), in *The Cochrane Library*, Vol. 1, 2001, Oxford: Update Software.
- [7] Cox, D. R., and Hinkley, D. V. (1974). *Theoretical Statistics*, London: Chapman and Hall.
- [8] Cramér. H. (1946). *Mathematical Method of Statistics*, Princeton: Princeton University Press.
- [9] Debboun, M., Strickman, D., Solberg, V. B., Wilkerson, R. C., McPherson, K. R., Golenda, C., Keep, L., Wirtz, R. A., Burge, R., and Klein, T. A. (2000). Field Evaluation of Deet and a Piperidine Repellent Against *Aedes Communis* (Diptera: Culicidae) and *Simulium Venustum* (Diptera: Simuliidae) in the Adirondack Mountains of New York, *Journal of Medical Entomology*, Vol. 37, 919-923.

- [10] Djordjevic, M. V., Stellman, S. D., and Zang, E. (2000). Doses of Nicotine and Lung Carcinogens Delivered to Cigarette Smokers, *Journal of the National Cancer Institute*, Vol. 92, 106-111.
- [11] Eising, E. G., and von der Ohe, M. R. (1998). Differentiation of Prolonged Colonic Transit Using Scintigraphy with Indium-111-Labeled Polystyrene Pellets, *Journal of Nuclear Medicine*, Vol. 39, 1062-1066.
- [12] Ezeamuzie, C. I., Al-Ali, S. F., Al-Dowaisan, A., Khan, M., Hijazi, Z., and Thomson, M.S. (1999). Reference Values of Total Serum IgE and Their Significance in the Diagnosis of Allergy Among the Young Adult Kuwaiti Population, *Clinical and experimental Allergy*, Vol. 29, 375-381.
- [13] Gotzsche, P. C. (2000). Do Patients with Osteoarthritis Get the Clinical Research They Need? *Annals of the Rheumatic Diseases*, Vol. 59, 407-408.
- [14] Herr, K. A., Mobily, P. R., Kohout, F. J., and Wagenaar, D. (1998). Evaluation of the Faces Pain Scale for Use with the Elderly, *The Clinical Journal of Pain*, Vol. 14, 29-38.
- [15] Inzelberg, R., Carasso, R. L., Schechtman, E., and Nispeanu, P. (2000). A Comparison of Dopamine Agonists and Catechol-O-Methyltransferase Inhibitors in Parkinson's Disease, *Clinical Neuropharmacology*, Vol. 23, 262-266.
- [16] Kann, L., Kinchen, S. A., Williams, B. I., Ross, J. G., Lowry, R., Grunbaum, J. A., and Kolbe, L. J. (2000). Youth Risk Behavior Surveillance—United States, 1999, *CDC Surveillance Summaries*, MMWR 2000, 49(SS-5).
- [17] Mancuso, C. A., Peterson, M. G. E., and Charlson, M. E. (2001). Comparing Discriminative Validity Between a Disease-Specific and a General Health Scale in Patients with Moderate Asthma, *Journal of Clinical Epidemiology*, Vol. 54, 263-274.
- [18] Martinez, M. E., Maltzman, T., Marshall, J. R., Einspahr, J., Reid, M. E., Sampliner, R., Ahnen, D. J., Hamilton, S. R., and Alberts, D. S. (1999). Risk Factors for Ki-Ras Protooncogene Mutation in Sporadic Colorectal Adenomas, *Cancer Research*, Vol. 59, 5181-5185.
- [19] McBride, N. (2000). The Western Australian School Health Project: Comparing the Effects of Intervention Intensity on Organizational Support for School Health Promotion, *Health Education Research*, Vol. 15, 59-72.
- [20] Mehta, S. R., Eikelboom, J. W., and Yusuf, S. (2000). Reply to comments on Intracranial Haemorrhage with Bolus Thrombolytic Agents, *The Lancet*, Vol. 356, 1850.
- [21] Millikan, R., Pittman, G., Tse, C. K., Savits, D. A., Newman, B., and Bell, D. (2000). Glutathione S-Transferases M1, T1, and P1 and Breast Cancer, *Cancer Epidemiology Biomarkers & Prevention*, Vol. 9, 567-573.
- [22] Rahlfs, V. W. (1997), Comment on Understanding and Evaluating Clinical Trials, *Journal of the American Academy of Dermatology*, Vol. 37, 803-804.
- [23] Schenker, N., and Gentleman, J. F. (2001). On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals, *American Statistician*, Vol. 55,

No. 3, 182-186.

- [24] Slavotinek, A., Rosenberg, M., Knight, S., Gaunt, L., Fergusson, W., Killoran, C., Clayton-Smith, J., Kingston, H., Campbell, R. H., Flint, J., Donnai, D., and Biesecker, L. (1999). Screening for Submicroscopic Chromosome Rearrangements in Children with Idiopathic Mental Retardation Using Microsatellite Markers for the Chromosome Telomeres, *Journal of Medical Genetics*, Vol. 36, 405-411.
- [25] Sont, W. N., Zielinski, J. M., Ashmore, J. P., Jiang, H., Krewski, D., Fair, M. E., Band, P. R., and Létourneau, E. G. (2001). First Analysis of Cancer Incidence and Occupational Radiation Exposure Based on the National Dose Registry of Canada, *American Journal of Epidemiology*, Vol. 153, 309-318.
- [26] Tersmette, A. C., Petersen, G. M., Offerhaus, G. J. A., Falatko, F. C., Brune, K. A., Goggins, M., Rozenblum, E., Wilentz, R. E., Yeo, C. J., Cameron, J. L., Kern, S. E., and Hruban, R. H. (2001). Increased Risk of Incident Pancreatic Cancer Among First-Degree Relatives of Patients with Familial Pancreatic Cancer, *Clinical Cancer Research*, Vol. 7, 738-744.
- [27] Vatassery, G. T., Smith, W. E., and Quach, H. T. (1998). α -Tocopherol in Rat Brain Subcellular Fractions is Oxidized Rapidly During Incubations with Low Concentrations of Peroxynitrite, *The Journal of Nutrition*, Vol. 128, 152-157.
- [28] Wong, C. M., Hu, Z. G., Lam, T. H., Hedley, A. J., and Peters, J. (1999). Effects of Ambient Air Pollution and Environmental Tobacco Smoke on Respiratory Health of Non-Smoking Women in Hong Kong, *International Journal of Epidemiology*, Vol. 28, 859-864.

[2002년 1월 접수, 2002년 5월 채택]