

A Comparison Study for the Confidence Intervals of the Common Odds Ratio in the Stratified 2×2 Tables Using the Average Coverage Probability¹⁾

Min Jung Kwak²⁾ and Hyeong Chul Jeong³⁾

Abstract

In this paper, various methods for finding confidence intervals for common odds ratio ψ of the K 2×2 tables are reviewed. Also we propose two jackknife confidence intervals and bootstrap confidence intervals for ψ . These confidence intervals are compared with the other existing confidence intervals by using Monte Carlo simulation with respect to the average coverage probability.

Keywords : Common odds ratio, Exact confidence interval, Jackknife, Bootstrap, Average coverage probability

1. 소개

사례대조연구나 임상시험 등의 자료 중 상당 부분은 K 개의 층(stratum)을 지닌 독립적인 2×2 이차원 분할표로 주어진다. 여기서, 층 변수의 효과를 통제하면서 두 범주형 변수 간의 동질연관성을 측정하는 방법 중 하나로 공통오즈비가 활용된다. 특히, 공통오즈비 추정과 더불어 공통오즈비 ψ 의 신뢰구간을 구축하는 문제는 많은 관심을 받아왔는데, 대표적으로 다음의 두 가지 방법을 들 수 있다. 첫째 방법은 Mantel-Haenszel 공통오즈비의 추정값과 이에 대한 근사적 분산추정치 를 결합하여 정규분포를 활용하여 신뢰구간을 구축하는 것이다. 이에 대한 연구로 Hauck(1979)의 신뢰구간, Phillips and Holland(1987) 그리고 Sato(1990)의 신뢰구간을 참고할 수 있다. 그런데 Robins, Breslow and Greenland(1986)의 정규근사 신뢰구간이 sparse 자료나 K 가 비교적 클 때도 신뢰구간을 잘 구축하는 것으로 알려져 있다. 두 번째는 초기하분포에 기초한 정확신뢰구간을 구축하는 것이다. 이에 대해서는 Mehta, Patel and Gray(1985), Vollset, Hirji and Elashoff(1991) 등의 연구를 들 수 있다. 정확신뢰구간은 sparse 자료나 치우침 정도가 심한 자료에도 신뢰구간을 구축할 수 있으나, 범주형자료의 이산성 때문에 포함확률이 주어진 명목수준보다 더 크게 되는 보수성 문제가 발생한다(Agresti and Coull; 1998). 정확신뢰구간의 보수성을 피하는 방법으로 중간

1) This research was supported by Pyongtaek University Research Grant in 2001.

2) Assistant Professor, Department of Information Statistics, Pyongtaek University, Kyunggi, 450-701, Korea E-Mail : mjkwak@ptuniv.ac.kr

3) Assistant Professor, Department of Information Statistics, Pyongtaek University, Kyunggi, 450-701, Korea E-Mail : jhc@ptuniv.ac.kr

-p 값(mid-p) 조정방법을 사용할 수 있다. 그런데, 이들 정확신뢰구간의 구축에는 많은 계산적 어려움이 수반된다고 하겠다.

공통오즈비의 신뢰구간 구축에는 위에서 언급한 Mantel-Haenszel 추정치와 초기하분포의 최대우도 추정치를 활용하는 방법 외에 재표본 설계(resampling plan)를 하는 잭나이프(Jackknife) 추정치와 붓스트랩 추정치를 활용할 수 있다. 특히, Pigeot(1991)는 층의 개수가 적을 때 각 셀에 해당되는 각각의 관측치들을 하나의 표본으로 간주하여 잭나이프를 하는 방법이 편향(bias)과 평균제곱오차(mean squared error) 측면에서 우수함을 보였다. 또한, 콰민정 등(2002)은 층의 수가 2일 때 붓스트랩 추정량이 편의와 평균제곱오차 측면에서 잭나이프 방법보다 우수함을 모의실험을 통해 보인바 있다. 그런데, Singh(1981)에 의하면, 초기하분포와 같이 격자분포를 따르는 분포에서 붓스트랩분포와 실제분포와의 차이가 $O(1/\sqrt{n})$ 인데, 정규근사를 이용한 분포와 실제분포와의 차이가 역시 $O(1/\sqrt{n})$ 으로 정규근사 방법과 붓스트랩 방법과의 차이가 없음을 보인바 있다. 하지만, Woodroffe and Jhun(1989)은 평균포함신뢰확률 측면에서 붓스트랩분포와 실제분포와의 차이를 실질적으로 $o(1/\sqrt{n})$ 까지 줄일 수 있음을 제시하였다. 이와같은 사실에 따라 공통오즈비의 신뢰구간 구축에도 붓스트랩 방법이 잘 적용되는지 살펴보는 것도 의미가 있으리라 여겨진다.

Mehta and Walsh(1992)는 여러 가지의 공통오즈비, 층의 수, 표본의 크기 등을 고려한 420가지의 경우 수에서 정확신뢰구간, 중간-p 조정 신뢰구간 그리고 Mantel-Haenszel 신뢰구간을 모의실험을 통해 비교한바 있는데, 본 연구에서는 기존의 여러 신뢰구간과 재표본 설계방법인 잭나이프와 붓스트랩 신뢰구간을 제시하고, 평균포함확률 측면에서 이들 신뢰구간과 추정량 그리고 평균길이 등을 비교하여 각 신뢰구간의 특징을 살펴보고자 한다. 2장에서는 공통오즈비 ϕ 에 대한 여러 신뢰구간을 살펴보고 3장에서는 평균포함확률 측면에서 모의실험을 통하여 이들을 비교하였다.

2. 공통오즈비 ϕ 의 신뢰구간 구축

독립적인 K개의 층화된 2x2 분할표가 주어질 때, $X_k=(X_{1k}, X_{2k})$ 를 k 층에서 크기가 각각 n_{1k}, n_{2k} 인 표본에서 성공횟수라 놓으면, $X_{1k} \sim B(n_{1k}, p_{1k})$, $X_{2k} \sim B(n_{2k}, p_{2k})$ 인 이항분포를 따르며, 각 층에서 사례-대조 연구의 2x2 분할표는 다음과 같이 생성된다.

	성공	실패	총합
사례	x_{1k}	$n_{1k} - x_{1k}$	n_{1k}
대조	x_{2k}	$n_{2k} - x_{2k}$	n_{2k}
총합	$t_k = x_{1k} + x_{2k}$	$n_k - t_k$	$n_k = n_{1k} + n_{2k}$

여기서, 오즈비는 $\phi_k = p_{1k}(1-p_{2k})/(1-p_{1k})p_{2k}$ 와 같다. 이제, 모든 층에서 오즈비가 동일하다는 가정, 즉 $\phi_1 = \phi_2 = \dots = \phi$ 하에서 공통오즈비의 신뢰구간 구축에는 근사 신뢰구간과 정확신뢰구간의 두 가지 접근방법이 가능하다. 본 절에서 각 접근방법에 대해 간략히 기술하고자 한다.

2.1 정확신뢰구간

S_k 를 층 k 에서 사례의 성공에 대한 확률변수라 놓으면(실제 관찰치는 x_{1k}), S_k 는 $\max(0, t_k - n_{2k})$ 와 $\min(n_{1k}, t_k)$ 에서 결정된다. 이제, $S = S_1 + S_2 + \dots + S_K$ 라 놓고, 실제 관찰치의 합을 $s = x_{11} + x_{12} + \dots + x_{1K}$ 라 놓으면, S 의 조건부 확률분포는 $T = (t_1, t_2, \dots, t_K)$ 가 주어질 때, 다음과 같다.

$$\Pr(S = s | T, \phi) = \frac{C_s \phi^s}{\sum_{x=s_{\min}}^{s_{\max}} C_x \phi^x}, \quad s_{\min} \leq s \leq s_{\max} \quad (2.1)$$

여기서, $s_{\min} = \sum_{k=1}^K \max(0, t_k - n_{2k})$, $s_{\max} = \sum_{k=1}^K \min(n_{1k}, t_k)$, 그리고 $C_s = \sum_{T(s)} \binom{n_{1k}}{s_k} \binom{n_{2k}}{t_k - s_k}$ 이며, $T(s) = [(s_1, \dots, s_K) : \max(0, t_k - n_{2k}) \leq s_k \leq \min(n_{1k}, t_k) \text{ and } s_1 + \dots + s_K = s]$ 이다.

2.1.1 정확신뢰구간

공통오즈비의 정확신뢰구간의 하한을 ψ_L , 상한을 ψ_U 라 놓으면, 식 (2.1)을 이용한 $100(1-\alpha)\%$ 정확신뢰구간은 다음과 같다.

$$\begin{aligned} \psi_L &= 0 & \text{if } s = s_{\min} & & \Pr(S = s_{\min} | \psi_U) &= \alpha \\ \psi_L: \Pr(S \geq s | \psi_L) &= \alpha/2 & \text{if } s_{\min} < s < s_{\max} & & \psi_U: \Pr(S \leq s | \psi_U) &= \alpha/2 & \text{if } s_{\min} < s < s_{\max} \\ \Pr(S = s_{\max} | \psi_L) &= \alpha & & & \psi_U &= \infty & \text{if } s = s_{\max} \end{aligned} \quad (2.2)$$

위의 정확신뢰구간을 유도하기 위해서는 먼저, 식 (2.1)에서 나타나는 계수 C_x 를 계산하여야 하며, 다음으로 식 (2.2)의 다항방정식을 풀어야 한다. 그런데, 표본의 크기가 조금만 커지더라도, 계수 C_x 의 값과 다항방정식에서 컴퓨터의 계산한계를 넘는 overflow 문제가 발생하므로 정확신뢰구간을 유도하기 위한 다양한 수치적 알고리즘이 연구되어 왔다. 계수 C_x 를 유도하는 대표적인 방법으로 Mehta, Patel and Gray(1985)의 network 알고리즘과 network 알고리즘을 개선한 Vollset, Hirji, and Elashoff(1991)의 recursive 알고리즘을 들 수 있다. 본 연구에서는 Hirji and Vollset(1993)의 정확분포를 유도하는 알고리즘 AS289를 사용하여, 계수 C_x 를 유도하였다. Hirji and Vollset(1993)의 AS289는 $\psi = 1$ 인 경우, 식 (2.1)의 확률을 계산하여 준다. 그러므로 정확분포의 확률에 모든 계수의 합에 해당하는 $\binom{n_k}{t_k}$ 를 곱하여 계수 C_x 를 유도할 수 있다. 다음으로 식 (2.2)의 다항식의 근을 구하기 위해서는 식 (2.2)를 다음과 같이 log 변환한다.

$$\log\left(\sum_{x=s}^{s_{\max}} c_x \phi^x\right) - \log\left(\sum_{x=s_{\min}}^{s_{\max}} c_x \phi^x\right) = \log(\alpha/2) \quad (2.3)$$

위의 식을 만족하는 ϕ 를 수치해석적 방법으로 유도하면 된다. 식 (2.3)의 $\log\left(\sum_{x=s}^{s_{\max}} c_x \phi^x\right)$ 를 유도

하기 위해서는 $a_i = \log(c_i \psi^i)$ 로 놓고, 두 로그값의 합을 유도하는 다음의 식 (2.4)를 통하여 overflow 문제를 극복할 수 있다(Mehta, and Patel and Gray; 1985).

$$c = \log(\exp(a) + \exp(b)) = \log(\exp(a-d) + \exp(b-d)) + d \quad (2.4)$$

여기서, $d = \max(a, b)$ 이다. Vollset, Hirji, and Elashoff(1991)는 식 (2.2)의 다항식에 대해 secant 방법을 활용하는 것이 bisection 방법 보다 효율적임을 제시한 바 있다.

2.1.2 중간-p 조정 신뢰구간

식 (2.2)의 정확신뢰구간은 자료의 이산성 때문에 주어진 명목신뢰수준을 과대추정하는 경향이 있다. 이를 보완하는 방법으로 중간-p 조정(mid-p adjustment)이 사용된다. 중간-p 조정방법을 통한 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\begin{aligned} \psi_L: \quad & \frac{1}{2} \Pr(S = s | \psi_L) + \Pr(X > s | \psi_L) = \alpha/2 \\ \psi_U: \quad & \frac{1}{2} \Pr(S = s | \psi_U) + \Pr(X < s | \psi_U) = \alpha/2 \end{aligned} \quad (2.5)$$

Mehta and Walsh(1992)는 모의실험을 통해 중간-p 조정 신뢰구간이 정확신뢰구간이 지니는 근본적인 보수성을 상당부분 완화시킴을 보인바 있다.

2.2 근사신뢰구간

공통오즈비를 추정하는 추정량과 그의 표준오차를 활용하여 정규근사를 통한 공통오즈비의 신뢰구간을 구축할 수 있다. 이의 대표적인 방법은 Mantel-Haenszel 추정량과 그의 표준오차를 활용하는 것이다. Mantel-Haenszel 추정량의 표준오차에 대해서는 Hauck(1979), Breslow(1981), Flanders(1985) 등의 연구가 있는데, Robins, Breslow and Greenland(1986)가 제안한 표준오차가 가장 효율적으로 알려져 있다.

2.2.1 Mantel-Haenszel 근사신뢰구간

$R_k = [x_{1k}(n_{2k} - x_{2k})] / n_k$, $S_k = [x_{2k}(n_{1k} - x_{1k})] / n_k$ 그리고 각 층에 대한 R_k, S_k 의 합을 각각 $R_+ = \sum_{k=1}^K R_k$, $S_+ = \sum_{k=1}^K S_k$ 라 놓자. $P_k = \frac{[x_{1k} + (n_{2k} - x_{2k})]}{n_k}$, $Q_k = \frac{[x_{2k} + (n_{1k} - x_{1k})]}{n_k}$

이면, Mantel-Haenszel 추정량은 $\hat{\psi} = R_+ / S_+$ 이며, Robins et al(1986)의 $\log(\hat{\psi})$ 의 분산은

$$\widehat{var}(\log(\hat{\psi})) = \frac{1}{2} \sum_{k=1}^K \left[\frac{P_k R_k}{R_+^2} + \frac{P_k S_k + Q_k R_k}{R_+ S_+} + \frac{Q_k S_k}{S_+^2} \right]$$

과 같다. 또한 $\widehat{var}(\hat{\psi}) = \widehat{var}(\log(\hat{\psi})) \times \hat{\psi}^2$ 이다. 그런데, $\widehat{var}(\log(\hat{\psi}))$ 를 활용하여 신뢰구간을 구축한 후 지수변환을 사용하는 방법이 정규분포에 더 근사하게 된다. 이에 따라 Robins, Breslow and Greenland(1986)의 $100(1-\alpha)\%$ 정규근사 신뢰구간의 하한과 상한 $[\psi_L, \psi_U]$ 은 다음과 같다.

$$[\exp(\log(\hat{\psi}) - Z_{\alpha/2} \widehat{se}(\log(\hat{\psi}))), \exp(\log(\hat{\psi}) + Z_{\alpha/2} \widehat{se}(\log(\hat{\psi})))] \quad (2.6)$$

2.2.2 Peto 근사신뢰구간

임상실험에서 Mantel-Haenszel 추정량의 변형적 방법으로 Peto 추정량이 많이 활용되고 있다 (Emerson; 1994). 이제, $E_k = \frac{t_k n_{1k}}{n_k}$, $V_k = \frac{t_k(n_k - t_k)n_{1k}n_{2k}}{n_k^2(n_k - 1)}$ 라 놓으면, K개의 분할표로 부

터 $A = \sum_{k=1}^K x_{1k}$, $E = \sum_{k=1}^K E_k$ 그리고 $V = \sum_{k=1}^K V_k$ 를 계산할 수 있다. 여기서, $P = (A - E)/V$ 는 $\log(\psi)$ 의 Peto 추정량이며, $1/V$ 는 분산 추정량이다. 그러므로 Peto 추정량을 통한 $100(1-\alpha)\%$ 근사신뢰구간의 하한과 상한 $[\psi_L, \psi_U]$ 은 다음과 같이 주어진다.

$$[\exp(P - Z_{\alpha/2} \sqrt{1/V}), \exp(P + Z_{\alpha/2} \sqrt{1/V})] \quad (2.7)$$

2.2.3 잭나이프 신뢰구간

공통오즈비에 대한 추정으로 잭나이프 방법을 생각할 수 있다. Breslow and Liang(1982)은 층을 제거하는 방법에 대한 잭나이프 추정량을 제시하였으며, Pigeot(1991)은 각 셀의 관찰값을 하나씩 제거하는 잭나이프 추정량을 제시하였다. 이제, 2.2.1 절의 Mantel-Haenszel 추정량을 $\hat{\psi}_{MH}$ 2x2 분할표의 셀을 각각 a, b, c, d 라 놓으면, 셀을 잭나이프 했을 때 의-값(pseudo-values)은 다음과 같이 주어진다.

$$J_{a,k} = N\hat{\psi}_{MH} - (N-1)\hat{\psi}_{a,k}, \quad N = \sum_{k=1}^K n_k$$

여기서, $\hat{\psi}_{a,k} = \frac{(x_{1k}-1)(n_{2k}-x_{2k})/(n_k-1) + \sum_{j=1, j \neq k}^K x_{1j}(n_{2j}-x_{2j})/n_j}{x_{2k}(n_{1k}-x_{1k})/(n_k-1) + \sum_{j=1, j \neq k}^K x_{2j}(n_j-x_{1j})/n_j}$ 이며, 같은 방법으로

$J_{b,k}, J_{c,k}, J_{d,k}$ 를 구할 수 있다. Pigeot(1991)의 잭나이프 공통오즈비 추정량과 분산은 다음과 같다.

$$\hat{J} = \frac{1}{N} \sum_{k=1}^K [x_{1k}J_{a,k} + (n_{1k}-x_{1k})J_{b,k} + x_{2k}J_{c,k} + (n_{2k}-x_{2k})J_{d,k}]$$

$$V(\hat{J}) = \frac{1}{N(N-1)} \sum_{k=1}^K [x_{1k}(J-J_{a,k})^2 + (n_{1k}-x_{1k})(J-J_{b,k})^2 + x_{2k}(J-J_{c,k})^2 + (n_{2k}-x_{2k})(J-J_{d,k})^2]$$

Pigeot(1991)은 잭나이프 추정량이 편향과 평균제곱오차가 다른 방법들에 비해 비교적 적은 수준임을 모의실험을 통하여 보인바 있다. 그런데, 잭나이프 방법은 셀의 값이 0인 경우에 계산상 근본적인 어려움이 존재함을 짐작할 수 있다.

이제, Pigeot(1991)의 잭나이프 추정량이 대표본 이론에 의해 정규분포로 수렴한다고 가정하고, 잭나이프 정규근사신뢰구간을 다음과 같이 제안하고자 한다.

$$[J - Z_{\alpha/2} se(\hat{J}), J + Z_{\alpha/2} se(\hat{J})] \quad (2.8)$$

그런데, 식 (2.8)의 신뢰구간은 음의 값을 포함할 수 있으므로, 로그오즈비를 이용한 잭나이프 신뢰구간을 다음과 같이 구축할 수 있다.

$$[\exp(\log(\hat{\eta}) - Z_{\alpha/2} \text{se}(\hat{\eta})/\hat{\eta}), \exp(\log(\hat{\eta}) + Z_{\alpha/2} \text{se}(\hat{\eta})/\hat{\eta})] \quad (2.9)$$

2.2.4 붓스트랩 신뢰구간

$\log(\psi)$ 에 대해, 각 층에서 각 셀의 관찰값을 B 회의 붓스트랩 반복을 통해 얻은 붓스트랩 추정량을 $\hat{\psi}^*$ 라고 하면, $\hat{\psi}^* = \sum_{b=1}^B \log(\hat{\psi}_b^*)/B$ 이며, 표준오차는 $\sqrt{\sum_{b=1}^B (\log(\hat{\psi}_b^*) - \hat{\psi}^*)^2 / B - 1}$ 로 주어진다. 여기서, $\log(\hat{\psi}_b^*)$ 는 붓스트랩 표본을 이용하여 구한 Mantel-Haenszel $\log(\psi)$ 추정치이다. 이제, 붓스트랩 추정량 자체가 대표본 이론에 의해 정규근사한다고 놓으면, 위의 붓스트랩 추정량과 표준오차로부터 정규분포를 활용한 붓스트랩 신뢰구간을 다음과 같이 구축할 수 있다.

$$[\exp(\hat{\psi}^* - Z_{\alpha/2} \text{se}(\hat{\psi}^*)), \exp(\hat{\psi}^* + Z_{\alpha/2} \text{se}(\hat{\psi}^*))] \quad (2.10)$$

그런데, 식 (2.10)의 정규근사를 통한 신뢰구간 대신, 붓스트랩 분포로부터 신뢰구간을 얻을 수 있다. 붓스트랩 분포를 통한 $100(1-\alpha)\%$ 붓스트랩 신뢰구간은 다음과 같이 제안할 수 있다.

$$[\exp(\log(\hat{\psi}_{MH}) - q_{(1-\alpha/2)} \text{se}(\log(\hat{\psi}_{MH}))), \exp(\log(\hat{\psi}_{MH}) + q_{(\alpha/2)} \text{se}(\log(\hat{\psi}_{MH})))] \quad (2.11)$$

여기서, $q_{(\alpha)}$ 는 다음의 B 회 붓스트랩 반복에 의해 얻은 $Z^*(b)$ 의 $100(1-\alpha)\%$ 분위수이며, $\hat{\psi}_{MH}$ 은 2.2.1 절의 Mantel-Hanenszel 추정량이다.

$$Z^*(b) = \frac{(\log(\hat{\psi}_b^*) - \log(\hat{\psi}_{MH}))}{\text{se}(\log(\hat{\psi}_{MH}))}, \quad b = 1, \dots, B$$

위의 식 (2.11) 붓스트랩 신뢰구간은 식 (2.6)의 신뢰구간에서 정규분포를 이용한 대신에 주축통계량(pivot statistics)을 통한 붓스트랩 분포를 이용한다는 점에 차이가 있다고 하겠다. 또한, 붓스트랩 분포를 활용하여 신뢰구간을 구축하면, 범주형자료와 같은 격자분포에 대해 Woodrooffe and Jhun(1989)의 평균포함확률측면에서 붓스트랩 분포의 수렴정도를 파악할 수 있다.

2.3 사례분석

[표 2.1]은 앞에서 언급한 8가지 방법에 대해 3가지 자료에서 95% 신뢰구간의 추정결과를 보여주고 있다. [표 2.1]에서 EXACT는 식 (2.2), MID-P는 식 (2.5), MH/RBG는 식 (2.6), PETO는 식 (2.7), JACK1은 식 (2.8), JACK2는 식 (2.9), BOOT1은 식 (2.10) 그리고 BOOT2는 식 (2.11)의 신뢰구간을 의미한다. 편의상 사례(II)와 사례(III)의 자료는 Emerson(1994)의 자료를 참고하였으며, Emerson(1994)은 대표본 설계 방법을 제외한 신뢰구간들의 특징을 해당 자료에서 훌륭히 서술하였으므로, 본 사례분석에서는 대표본 설계 신뢰구간이 지니는 특징을 간략히 언급하고자 한다.

[표 2.1]에 제시된 중심값은 해당 신뢰구간의 로그공통오즈비의 중심을 계산한 후 다시 지수변환하여 얻은 값이다. 이러한 중앙값은 MID-P와 BOOT2를 제외한 나머지 방법들에서 공통오즈비의 추정값을 의미한다고 할 수 있다. 여기서, 중간-p 조정 방법과 붓스트랩 방법(BOOT2)은 로그오즈비를 중심으로 대칭의 분포를 따르지 않는다. 또한 이 두 방법으로 공통오즈비를 추정하는 것

[표 2.1] 층이 5일 때, 8가지 방법에 대한 공통오즈비의 신뢰구간과 추정치

k	사례(I)					사례(II)					사례(III)				
	a	b	c	d	odds	a	b	c	d	odds	a	b	c	d	odds
1	30	30	30	30	1	2	8	10	80	2.0	1	0	0	5	-
2	30	30	30	30	1	4	6	15	75	3.3	0	1	5	0	0
3	30	30	30	30	1	3	12	5	80	4.0	5	0	0	1	-
4	30	30	30	30	1	5	10	10	75	3.8	0	5	1	1	0
5	30	30	30	30	1	6	9	5	80	10.7	1	1	1	5	5
방법	중심값		95% 신뢰구간			중심값		95% 신뢰구간			중심값		95% 신뢰구간		
EXACT	0.9999		[0.7176, 1.3934]			4.0659		[2.0808, 7.9450]			1.8129		[0.1573, 20.8963]		
MID-P	1.0000		[0.7265, 1.3765]			4.0731		[2.1761, 7.6239]			1.8925		[0.2288, 15.6549]		
MH/RBG	1.0000		[0.7261, 1.3772]			4.1713		[2.2376, 7.7762]			1.9518		[0.5545, 6.8706]		
PETO	1.0000		[0.7271, 1.3754]			6.6873		[3.0456, 14.6835]			1.5158		[0.3578, 6.4212]		
JACK1	0.9797		[0.6572, 1.3022]			3.9233		[1.2723, 6.5744]			1.5941		[-1.4912, 4.6795]		
JACK2	0.9797		[0.7049, 1.3616]			3.9233		[1.9962, 7.7111]			1.5941		[0.2301, 11.0428]		
BOOT1	1.0073		[0.7320, 1.3861]			4.2460		[2.2512, 8.0081]			1.9721		[1.1319, 3.4362]		
BOOT2	0.9931		[0.7234, 1.3634]			4.0778		[2.1735, 7.6505]			1.9259		[1.1440, 3.2420]		

은 상당한 이견이 있을 수 있으나, 신뢰구간의 중심값을 이 두 방법의 공통오즈비의 추정치로 사용하는 것에는 큰 무리는 없으리라 여겨진다.

이제, 각 사례에서 나타난 신뢰구간의 특징을 살펴보자. [표 2.1]에서 사례(I)는 각 층에서 표본의 크기가 각각 120이며, 각 층의 오즈비가 1로 서로 동질적인 경우라 하겠다. 이때 대부분의 신뢰구간은 비교적 공통오즈를 1로 추정하나, 로그오즈를 사용하지 않은 잭나이프 신뢰구간은 하향추정의 경향을 보여주고 있다. 또한 Mantel-Haenszel 신뢰구간과 중간-p 조정신뢰구간, 그리고 Peto 신뢰구간이 상당히 유사하게 형성됨을 볼 수 있다.

사례(II)는 각 층의 표본의 크기가 각각 100으로 표본의 크기는 충분히 크며, 각 층간이 상당히 동질적이거나, 사례-대조 간에 표본의 크기가 이질적인 경우이다. 여기서, 정확신뢰구간의 최대우도 추정값을 최적이라고 놓으면, EXACT는 4.0659로 추정하나, PETO는 6.6873으로 Peto 신뢰구간 역시 상당히 위로 추정되고 있으며(Emerson; 1994), 잭나이프 신뢰구간은 하향추정 되고 있음을 볼 수 있다. 여기서, 정규근사를 사용한 붓스트랩 신뢰구간(BOOT1)은 정확방법에 비해 상향추정의 경향을 보이고 있다. 그런데, 주축통계량을 사용한 붓스트랩 신뢰구간(BOOT2)은 중간-p조정신뢰구간과 비슷한 결과를 유도함을 볼 수 있다. 즉 표본의 크기가 적당히 크면, 사례-대조 표본의 크기가 이질적이더라도, 붓스트랩 분포를 사용한 신뢰구간(BOOT2)이 정규근사를 통한 붓스트랩 신뢰구간(BOOT1) 보다 더 안정적인 결과를 유도한다고 하겠다.

사례(III)는 각 층간에 오즈비가 상당히 이질적이며, 표본의 크기 역시 상당히 작은 sparse 한 경우라 하겠다. 여기서, 계산의 편의를 위해서 정확신뢰구간인 경우를 제외하고 각 칸에 0.5 씩 더 하여 신뢰구간을 계산하였다. 이제, 정확신뢰구간의 최대우도추정값을 최적이라고 놓으면, Peto 추정치와 잭나이프 추정치는 하향추정, 붓스트랩 추정치와 Mantel-Haenszel 추정치는 상향추정 되고 있음을 볼 수 있다. 그런데, 신뢰구간 측면에서 JACK2 방법은 정확신뢰구간처럼 상향의 값을 비교적 높이 추정해 주며, 붓스트랩 신뢰구간은 상당히 좁은 폭의 신뢰구간을 형성하고 있음을 볼 수 있다. 이는 표본의 크기가 작으므로, 붓스트랩 분포가 Mantel-Haenszel 추정치를 중심으로 집

중되고 있기 때문이다. 즉 표본의 크기가 작을 때는 붓스트랩 추정치는 MID-P와 비슷하나, 신뢰구간의 폭이 상당히 좁아지는 경향을 보인다고 하겠다. 본 사례에서는 붓스트랩 분포의 분위점이 정규분포의 95%에 해당되는 분위점 1.96에 비해 상당히 낮은 수준에서 결정되었다.

3. 모의실험 비교

앞의 사례분석은 자료의 특수한 경우에 대해 신뢰구간의 특징을 살펴본 것이다. 이제, 자료의 특징을 보다 일반화 시켜 평균포함확률 측면에서 신뢰구간의 특징을 살펴보도록 하겠다. 평균포함확률이란 공통오즈비 ψ 에 대해, 어떤 사전분포를 가정하여, 실제모수 ψ 가 어떤 값으로 주어지던, 주어진 신뢰구간의 평균적 수행능력을 측정하는 것이라 하겠다. 이항분포나 포아송분포에서 평균포함확률의 적용은 Kim and Jeong(2000a, 2000b)을 참고할 수 있다.

3.1 모의실험과정

공통오즈비 ψ 의 평균포함확률을 추정하기 위해서, 고려해야할 변수가 $\psi, p_{1k}, n_{1k}, n_{2k}$ 그리고 K 등이 존재한다. 본 모의실험에서는 (1) $\log(\psi) \sim U[0, 3]$ (2) $\log(\psi) \sim |N(0, 1)|$ 를 가정하였다. 또한, 대조군의 확률 p_{2k} 에 대해서는 $U[0.2, 0.4]$ 를 가정하였다. 따라서 사례군 확률 p_{1k} 는 $\psi p_{2k} / (1 - p_{2k} + \psi p_{2k})$ 로 주어진다. 표본의 크기 $n = \sum_{k=1}^K n_k$ 은 20, 40, 80, 160 그리고 320을 고려하였으며, 대조군과 사례군의 표본의 비율은 각각 1:1과 1:3인 경우를 다루었다. 이에 따라 층의 수 K 를 2와 5로 제한하였다. 명목수준 $1 - \alpha$ 는 0.9, 0.95 그리고 0.99를 상정하였다. 모의실험 과정을 요약하면 다음과 같다.

[단계 1] $\log(\psi)$ 의 분포로부터 모수 ψ_i 를 랜덤추출한다.

[단계 2] $(p_{2k}; k=1, \dots, K)$ 의 분포에서 모수 $(p_{2k}; k=1, \dots, K)$ 를 랜덤추출하고, 공통오즈비를 유지하는 모수 $(p_{1k}; k=1, \dots, K)$ 를 계산한다.

[단계 3] $(p_{1k}, p_{2k}; k=1, \dots, K)$ 로부터 독립적으로 이항분포의 확률난수를 생성하여, $K \times 2$ 분할표를 생성한다. 주어진 분할표로부터 각 방법들의 신뢰구간을 계산한 후 $I(\psi_i)$, $L(\psi_i)$, $LC(\psi_i)$ 를 계산한다. 여기서, $I(\psi_i)$ 는 신뢰구간이 모수 ψ_i 를 포함하면 값이 1인 지시함수, $L(\psi_i)$ 는 신뢰구간의 길이, $LC(\psi_i)$ 는 신뢰구간의 중심값이다.

[단계 4] 위의 [단계1]~[단계3]을 M 회 반복한다.

[단계 5] 다음의 통계량을 계산한다.

$$\bar{\beta} = \sum_{i=1}^M I(\psi_i) / M, \quad \bar{L} = \sum_{i=1}^M L(\psi_i) / M, \quad r = \sum_{i=1}^M I_{[LC(\psi_i) \geq \psi_i]} / M.$$

[단계 5]에서 $\bar{\beta}$ 는 평균포함확률, \bar{L} 은 신뢰구간의 평균길이, 그리고 r 은 공통오즈비의 추정값

이 모수 ψ 에 비해 위에 존재하는 정도를 측정하는 통계량이다. (2.3)절에서 언급한 바와 같이 $\hat{\psi}_i = LC(\psi_i)$ 이므로 ψ 에 대한 추정치가 특별한 치우침이 없다면 r 은 0.5에 근사하게 될 것이다.

3.2 모의실험 결과

[표 3.1]과 [표 3.2]는 사전분포가 균등분포일 때 층의 수 2와 5에서 다양한 표본의 크기에 대해, 평균포함확률과 신뢰구간의 평균길이 그리고 r 을 명목수준 90%, 95% 그리고 99%에서 계산한 결

[표 3.1] 로그오즈비의 사전분포가 균등분포를 따를 때 각 방법들의 평균포함확률 (K=2, 사례군과 대조군의 비 1:1, 1:3, 모의실험횟수 1000회, 붓스트랩 반복 1000회)

N	방법	N1:N2 (1:1)						N1:N2 (1:3)											
		0.90	길이	R	0.95	길이	R	0.99	길이	R	0.90	길이	R	0.95	길이	R			
20	EXACT	.950	-	.576	.981	-	.596	.999	-	.633	.955	-	.577	.982	-	.591	.999	-	.609
	MID-P	.891	-	.557	.952	-	.578	.995	-	.617	.900	-	.563	.956	-	.577	.994	-	.606
	MH/RBG	.932	42.7	.452	.964	61.5	.452	.993	124.5	.452	.940	46.2	.439	.971	67.9	.439	.998	142.8	.439
	PETO	.917	16.5	.282	.955	22.1	.282	.991	38.0	.282	.925	23.8	.357	.964	32.9	.357	.995	60.8	.357
	JACK1	.777	27.5	.408	.856	32.8	.408	.927	43.1	.408	.739	29.6	.468	.830	35.3	.468	.921	46.4	.468
	JACK2	.762	-	.408	.781	-	.408	.870	-	.408	.666	73.3	.468	.747	145.4	.468	.864	752.2	.468
	BOOT1	.846	23.6	.483	.910	31.0	.483	.966	51.3	.483	.803	21.0	.461	.875	27.5	.461	.933	45.0	.461
	BOOT2	.816	23.6	.505	.856	33.3	.526	.905	64.0	.584	.755	19.5	.447	.823	26.6	.460	.880	51.0	.492
40	EXACT	.950	-	.545	.975	-	.557	.992	-	.585	.950	-	.537	.980	-	.549	.997	-	.576
	MID-P	.913	-	.532	.958	-	.546	.985	-	.573	.911	-	.525	.956	-	.537	.994	-	.560
	MH/RBG	.927	29.6	.535	.966	39.7	.535	.993	68.9	.535	.933	32.9	.512	.970	45.1	.512	.996	81.9	.512
	PETO	.922	13.1	.388	.961	16.7	.388	.992	25.9	.388	.926	19.8	.450	.963	26.0	.450	.995	42.8	.450
	JACK1	.818	20.3	.230	.859	24.2	.230	.911	31.8	.230	.799	23.3	.325	.835	27.7	.325	.892	36.4	.325
	JACK2	.937	-	.230	.966	-	.230	.985	-	.230	.927	85.3	.325	.948	196.8	.325	.961	-	.325
	BOOT1	.891	23.5	.566	.935	30.4	.566	.981	48.6	.566	.874	22.2	.529	.927	28.8	.529	.972	46.2	.529
	BOOT2	.834	24.9	.575	.875	33.6	.582	.936	60.3	.621	.808	24.0	.550	.861	32.9	.555	.923	59.0	.590
80	EXACT	.938	19.1	.531	.974	24.7	.537	.993	40.1	.553	.937	334.1	.510	.960	352.6	.518	.999	459.4	.540
	MID-P	.909	15.7	.529	.953	20.6	.532	.992	33.7	.547	.904	324.6	.501	.952	320.2	.512	.992	399.0	.534
	MH/RBG	.905	15.3	.530	.954	19.4	.530	.992	29.2	.530	.904	19.2	.512	.958	24.9	.512	.999	39.8	.512
	PETO	.907	8.7	.375	.950	10.8	.375	.993	15.5	.375	.901	12.7	.468	.951	16.0	.468	.995	23.9	.468
	JACK1	.861	14.8	.307	.894	17.7	.307	.946	23.3	.307	.854	16.5	.287	.889	19.7	.287	.932	25.9	.287
	JACK2	.955	21.8	.307	.980	30.5	.307	.994	58.8	.307	.943	26.5	.287	.966	38.3	.287	.988	80.5	.287
	BOOT1	.895	16.4	.567	.948	20.7	.567	.988	31.1	.567	.889	17.3	.545	.938	22.0	.545	.985	33.4	.545
	BOOT2	.868	14.5	.530	.918	18.6	.537	.967	29.4	.578	.850	17.3	.515	.895	22.5	.529	.944	37.1	.576
160	EXACT	.947	9.5	.538	.975	11.6	.544	.996	16.4	.559	.926	15.2	.523	.962	20.7	.528	.990	46.2	.545
	MID-P	.917	8.6	.538	.963	10.6	.540	.993	15.2	.552	.894	12.4	.519	.947	16.6	.523	.988	22.8	.536
	MH/RBG	.916	8.6	.546	.963	10.5	.546	.994	14.7	.546	.888	11.3	.525	.950	14.0	.525	.990	20.2	.525
	PETO	.902	5.8	.359	.959	7.1	.359	.993	9.7	.359	.889	8.4	.511	.942	10.3	.511	.986	14.5	.511
	JACK1	.893	8.5	.430	.927	10.1	.430	.971	13.3	.430	.877	11.1	.405	.903	13.2	.405	.964	17.4	.405
	JACK2	.954	9.3	.430	.981	11.5	.430	.995	16.5	.430	.944	13.5	.405	.981	17.6	.405	.996	28.8	.405
	BOOT1	.908	9.3	.577	.962	11.4	.577	.995	16.0	.577	.881	11.9	.556	.937	14.7	.556	.986	21.1	.556
	BOOT2	.906	8.3	.541	.952	10.2	.543	.986	14.7	.577	.861	10.8	.525	.922	13.4	.539	.967	20.1	.571
320	EXACT	.924	5.9	.523	.962	7.0	.531	.993	9.5	.545	.930	7.0	.497	.971	8.5	.498	.997	11.6	.509
	MID-P	.897	5.4	.516	.951	6.6	.525	.991	9.1	.541	.911	6.4	.496	.962	7.8	.498	.994	10.9	.505
	MH/RBG	.897	5.5	.535	.950	6.6	.535	.991	9.0	.535	.906	6.4	.507	.961	7.8	.507	.995	10.7	.507
	PETO	.858	3.9	.307	.923	4.7	.307	.980	6.4	.307	.900	5.5	.506	.957	6.7	.506	.993	9.1	.506
	JACK1	.907	5.4	.454	.940	6.5	.454	.971	8.5	.454	.897	6.4	.420	.940	7.6	.420	.979	10.0	.420
	JACK2	.928	5.6	.454	.964	6.8	.454	.995	9.3	.454	.941	6.7	.420	.977	8.2	.420	.997	11.4	.420
	BOOT1	.890	5.7	.562	.946	6.8	.562	.989	9.3	.562	.903	6.7	.534	.963	8.2	.534	.996	11.2	.534
	BOOT2	.891	5.4	.527	.948	6.5	.535	.985	9.0	.562	.896	6.3	.499	.956	7.7	.507	.987	10.7	.549

※ 각 방법들은 앞에서 제시되었으며, 신뢰구간의 평균길이가 - 인 부분은 모의실험과정 중 발생된 랜덤표본에 대해 신뢰구간의 값이 무한대가 되거나 신뢰구간의 길이가 너무 길어 전체적인 특징을 평균으로 나타낼 수 없는 경우가 있음을 나타낸다. 신뢰구간을 유도할 수 없는 표본에 대해서는 연속성 수정을 가하였다.

과를 보여주고 있다.

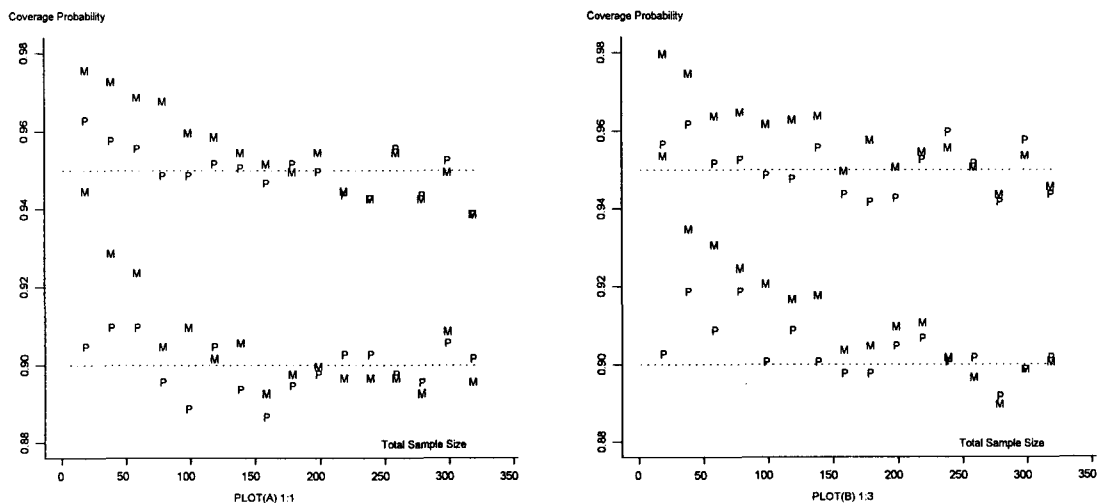
[표 3.1]의 결과를 살펴보면, 식 (2.8)의 ϕ 에 대한 잭나이프 신뢰구간(JACK1)이 가장 하향추정하고 있음을 볼 수 있다. 그런데, 이를 $\log(\phi)$ 로 변환하여 유도한 잭나이프 신뢰구간(JACK2)은 JACK1을 상당부분 조정하여 주나 표본의 크기가 커짐에 따라 정확신뢰구간의 명목수준을 넘어서는 과대추정의 경향을 보이고 있다.

[표 3.2] 로그오즈비의 사전분포가 균등분포를 따를 때 각 방법들의 평균포함확률 (K=5, 사례군과 대조군의 비 1:1, 1:3, 모의실험횟수 1000회, 붓스트랩 반복 1000회)

N	방법	N1:N2 (1:1)						N1:N2 (1:3)											
		0.90	길이	R	0.95	길이	R	0.99	길이	R	0.90	길이	R	0.95	길이	R	0.99	길이	R
20	EXACT	.964	-	.635	.983	-	.656	.995	-	.689	.974	-	.620	.992	-	.649	.998	-	.684
	MID-P	.905	-	.610	.968	-	.640	.991	-	.679	.940	-	.575	.976	-	.619	.996	-	.673
	MH/RBG	.875	15.3	.238	.942	20.6	.238	.986	35.5	.238	.882	14.7	.221	.945	19.8	.221	.988	34.7	.221
	PETO	.774	8.2	.095	.883	10.7	.095	.970	17.1	.095	.807	9.3	.129	.891	12.3	.129	.979	20.3	.129
	JACK1	.616	10.6	.357	.698	12.6	.357	.808	16.6	.357	.608	9.8	.350	.700	11.7	.350	.800	15.4	.350
	JACK2	.782	13.8	.357	.846	18.3	.357	.881	31.6	.357	.782	14.1	.350	.837	19.9	.350	.884	41.5	.350
	BOOT1	.621	5.9	.245	.705	7.3	.245	.851	10.4	.245	.476	4.4	.219	.568	5.4	.219	.702	7.5	.219
	BOOT2	.556	5.5	.247	.628	6.7	.250	.765	10.0	.276	.444	4.4	.228	.524	5.3	.228	.671	7.9	.264
40	EXACT	.932	-	.549	.960	-	.563	.996	-	.603	.943	-	.561	.969	-	.578	.992	-	.615
	MID-P	.900	-	.531	.939	-	.550	.988	-	.585	.896	-	.540	.947	-	.561	.987	-	.600
	MH/RBG	.907	13.8	.318	.946	17.8	.318	.990	28.4	.318	.903	14.2	.324	.951	18.5	.324	.988	30.0	.324
	PETO	.832	7.8	.168	.909	9.9	.168	.977	14.9	.168	.863	9.7	.219	.924	12.4	.219	.983	19.2	.219
	JACK1	.729	12.0	.253	.795	14.3	.253	.867	18.9	.253	.697	11.7	.305	.761	14.0	.305	.867	18.4	.305
	JACK2	.907	16.1	.253	.941	21.6	.253	.962	37.4	.253	.890	16.1	.305	.925	21.9	.305	.937	39.7	.305
	BOOT1	.803	9.2	.350	.875	11.4	.350	.949	16.6	.350	.719	7.9	.347	.817	9.8	.347	.923	14.1	.347
	BOOT2	.740	8.4	.320	.818	10.5	.327	.919	16.0	.359	.679	7.3	.323	.757	9.1	.331	.879	13.5	.350
80	EXACT	.926	129.6	.507	.966	144.9	.518	.988	232.4	.550	.936	646.2	.522	.965	677.9	.531	.991	891.1	.553
	MID-P	.896	121.4	.500	.941	131.5	.508	.984	182.5	.536	.903	629.8	.519	.948	648.4	.524	.987	763.6	.545
	MH/RBG	.925	11.6	.417	.960	14.5	.417	.991	21.4	.417	.922	11.9	.390	.965	15.0	.390	.991	22.7	.390
	PETO	.881	7.0	.240	.942	8.7	.240	.983	12.4	.240	.904	8.9	.325	.943	11.2	.325	.986	16.4	.325
	JACK1	.845	11.1	.297	.895	13.2	.297	.943	17.4	.297	.804	10.8	.314	.865	12.9	.314	.921	16.9	.314
	JACK2	.933	13.4	.297	.961	17.2	.297	.991	27.0	.297	.916	13.1	.314	.953	16.9	.314	.983	26.8	.314
	BOOT1	.875	9.4	.418	.927	11.6	.418	.976	16.7	.418	.836	8.8	.415	.905	10.9	.415	.967	15.6	.415
	BOOT2	.842	9.8	.453	.903	12.2	.456	.958	18.2	.506	.814	8.9	.415	.871	11.0	.418	.953	16.4	.457
160	EXACT	.919	9.9	.501	.961	12.1	.509	.987	17.2	.528	.930	13.6	.503	.956	17.2	.510	.987	27.1	.529
	MID-P	.895	8.8	.500	.944	10.9	.503	.983	15.8	.522	.907	11.6	.496	.942	14.9	.502	.980	23.6	.522
	MH/RBG	.896	8.7	.513	.948	10.7	.513	.986	15.0	.513	.918	9.4	.479	.953	11.6	.479	.989	16.5	.479
	PETO	.891	5.6	.322	.941	6.8	.322	.989	9.4	.322	.911	7.5	.432	.949	9.2	.432	.989	12.9	.432
	JACK1	.889	8.6	.387	.923	10.3	.387	.962	13.6	.387	.879	9.1	.378	.921	10.9	.378	.960	14.3	.378
	JACK2	.928	9.5	.387	.967	11.7	.387	.988	16.9	.387	.931	10.1	.378	.962	12.6	.378	.990	18.5	.378
	BOOT1	.891	8.0	.522	.945	9.7	.522	.988	13.5	.522	.890	8.0	.477	.935	9.8	.477	.980	13.7	.477
	BOOT2	.846	8.4	.519	.903	10.3	.527	.958	14.8	.556	.850	8.4	.500	.912	10.4	.512	.959	14.9	.540
320	EXACT	.939	5.7	.520	.966	6.9	.524	.994	9.3	.538	.924	7.1	.529	.962	8.6	.534	.996	11.8	.547
	MID-P	.910	5.3	.518	.958	6.5	.521	.993	8.9	.537	.904	6.5	.527	.948	8.0	.532	.990	11.1	.545
	MH/RBG	.911	5.4	.534	.957	6.5	.534	.993	8.9	.534	.910	6.5	.550	.954	7.9	.550	.990	10.9	.550
	PETO	.861	3.9	.300	.926	4.6	.300	.981	6.3	.300	.908	5.4	.496	.943	6.6	.496	.989	9.0	.496
	JACK1	.914	5.4	.431	.940	6.5	.431	.972	8.5	.431	.906	6.5	.441	.944	7.8	.441	.974	10.2	.441
	JACK2	.940	5.6	.431	.968	6.8	.431	.997	9.3	.431	.935	6.9	.441	.968	8.4	.441	.995	11.6	.441
	BOOT1	.904	5.6	.577	.954	6.7	.577	.993	9.1	.577	.904	6.3	.579	.949	7.6	.579	.987	10.4	.579
	BOOT2	.896	5.2	.521	.949	6.3	.526	.989	8.7	.564	.897	6.2	.545	.944	7.5	.555	.988	10.6	.590

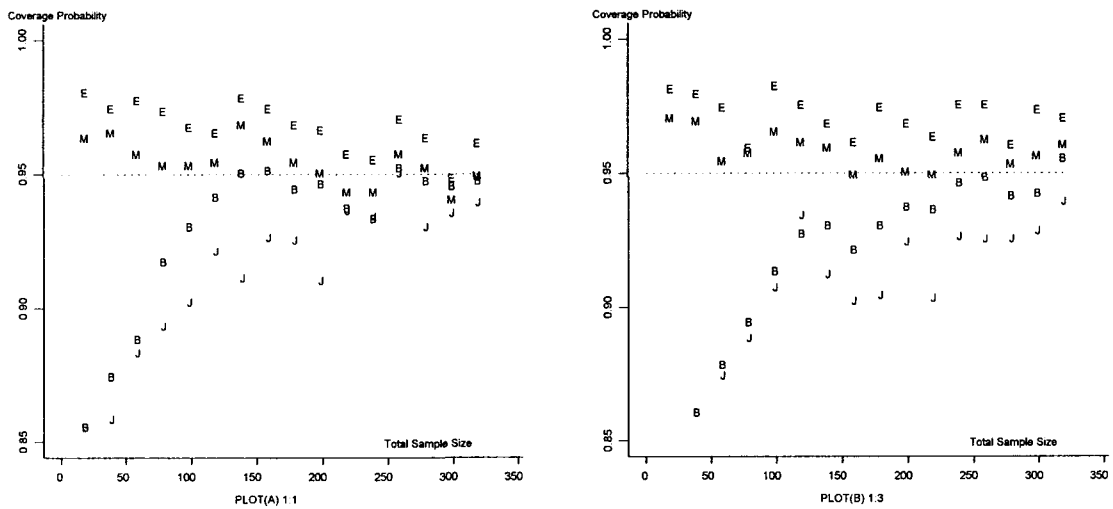
먼저, 모의실험결과를 통하여, 정확신뢰구간과 공통오즈비의 대표적 신뢰구간으로 알려진 MH/RBG 신뢰구간을 비교하여 보자. 정확신뢰구간(EXACT)은 [표 3.1], [표 3.2]에서 모두 주어진 명목수준보다 과대추정 되고 있으며, 이는 정확신뢰구간이 지니는 보수적 성향을 보여준다고 하겠

다. 이러한 결과는 [그림 3.2]에서 재확인할 수 있다. 이에 반해 중간-p 조정 신뢰구간은 정확신뢰구간의 보수성을 상당부분 완화시켜주며, 주어진 명목수준에도 잘 근사하는 방법임을 알 수 있다. 그리고 MH/RBG 방법은 표본의 크기가 작을 때, 평균포함확률 측면에서 MID-P에 비해 다소 과대 추정되는 경향이 있으나, 표본의 크기가 커지면서 MID-P 방법에 비해 덜 보수적인 경향을 보이고 있음을 알 수 있다. 또한 표본의 크기가 커짐에 따라 신뢰구간의 폭이 MH/RBG 방법이 MID-P에 비해 작음을 알 수 있다. MID-P와 MH/RBG의 경향을 좀더 살펴보자. [그림 3.1]은 $K = 5$ 에서, MID-P 방법과 MH/RBG 방법의 90%, 95% 명목수준에서 $\log(\phi)$ 의 분포가 정규분포일 때 평균포함확률을 보여준다. [그림 3.1]은 평균포함확률 측면에서 상당히 흥미로운 결과를 보여주고 있다. [그림 3.1]에서 표본의 크기가 작을 때 두 방법 모두 주어진 명목수준 90%, 95%를 넘는 과대추정 경향이 있으나, 표본의 크기가 커짐에 따라 주어진 명목수준에 평균포함확률이 근접해감을 볼 수 있다. 특히, MH/RBG 방법은 표본의 크기가 작을 때 대부분 MID-P 평균포함확률 보다 높은 수준에서 결정되고 있음을 볼 수 있다. 이러한 경향은 사례-대조의 비가 1:1에서 1:3으로 증가할수록 명백해진다. 이는 표본의 크기가 작을 때, MH/RBG 방법은 연속성 조정을 해야 할 필요성이 발생하기 때문으로 짐작할 수 있다. 대부분의 연구에서 MH/RBG 방법은 여러 가지 표본 상황에서 주어진 명목 수준을 잘 추정하면서도 계산이 정확방법에 비해 상당히 간편하기에 많이 선호되어 왔다. 이에 대해, Mehta and Walsh(1992)는 MH/RBG 방법은 신뢰구간을 계산할 수 없는 경우가 있으므로, 대부분의 적절한 자료에서 MH/RBG를 활용할 수 있으나, 그렇지 않은 경우에는 EXACT 신뢰구간을 사용하는 것이 효율적이라 지적하였다. 그런데, 평균포함확률 측면에서 보면, MH/RBG 방법은 표본의 크기가 작을 때는 MID-P 방법보다 다소 보수적 성향을 보여주므로, 표본의 크기가 작을 때는 MID-P 방법이 MH/RBG 방법보다 더 우수하며, 이는 사례-대조의 비가 이질적 경향이 커질수록 MID-P 방법이 MH/RBG 보다 더 선호될 수 있음을 보여준다고 하겠다.



[그림 3.1] $\log(\phi) \sim N[0, 1]$ 일 때, 층의 수가 5에서 명목수준 90%, 95%에 대한 MH/RBG와 MID-P 평균포함확률(M : MH/RBG, P : MID-P)
[PLOT(A)] 사례-대조 비(1:1), [PLOT(B)] 사례-대조 비 (1:3)

모의실험 결과, PETO 신뢰구간은 표본의 크기가 증가함에 따라 주어진 명목수준을 하향추정해 가는 경향을 보여준다. 하향추정의 결과로 PETO 신뢰구간은 신뢰구간의 평균 폭이 다른 방법에 비해 좁음을 볼 수 있다. 그런데, PETO 신뢰구간은 결정적인 약점을 지니고 있음을 r 에서 찾을 수 있다. PETO 신뢰구간의 r 은 거의 모든 경우에 항상 0.5보다 작으며, 이는 총의 수가 큰 [표 3.2]에서 그 경향이 더 심해지고 있음을 볼 수 있다. 이는 PETO 추정량이 대부분의 경우 모수 ϕ 를 하향추정하고 있어 편향(bias)이 큰 통계량임을 의미하며, 이에 따라 신뢰구간이 다른 방법에 비해 평균적으로 아래에서 좁게 형성되고 있음을 나타내는 결과라 하겠다.



[그림 3.2] $\log(\phi) \sim U[0, 3]$ 일 때, 총의 수가 2에서 표본의 크기에 따른 평균포함확률 (E : EXACT, M : MH/RBG, B : BOOT2, J : JACK1) [PLOT(A)] 사례-대조 비(1:1), [PLOT(B)] 사례-대조 비 (1:3)

잭나이프 신뢰구간은 대부분 주어진 ϕ 를 하향추정하고 있음을 볼 수 있다. 또한, Pigeot(1991)의 ϕ 에 대한 평균과 분산을 통해 잭나이프 신뢰구간을 구축하는 것보다, $\log(\phi)$ 변환으로 신뢰구간을 구축하는 것(JACK2)이 좀더 효율적인 방법임을 보여 준다고 하겠다. 그런데, JACK2 신뢰구간은 [표 3.1]과 [표 3.2]를 통하여, 표본의 크기가 작을 때는 심각한 하향추정의 경향을 보이나, 표본의 크기가 조금 커짐에 따라 JACK1의 신뢰구간을 상당히 보정하여 주는 방법임을 볼 수 있다. 그런데, JACK2는 표본의 크기가 커짐에 따라 평균포함률이 EXACT 평균포함확률과 비슷한 경향을 보이고 있으며, 잭나이프 추정량 역시 PETO 추정량처럼 주어진 모수 ϕ 를 대부분 밑에서 추정하여 r 이 낮게 형성되고 있다.

이제, 평균포함확률 측면에서 붓스트랩 방법이 지니는 의미를 보다 정확히 파악하여 보자. [그림 3.2]는 $K=2$ 에서 EXACT, MH/RBG, BOOT2 그리고 JACK1의 평균포함확률을 보여주고 있다. [그림 3.2]에서, 붓스트랩 주축 통계량의 분포를 활용한 붓스트랩 신뢰구간의 수립하는 형태를 볼 수 있다. 사례대조의 비가 1:1일 때는 총 표본의 크기가 약 100 이상에서, 1:3 일 때는 총 표본의

크기가 약 200 이상에서, 주어진 명목수준에 수렴하고 있음을 보여 준다. 특히, $K=5$ 등 층의 수가 증가할 수록 붓스트랩 신뢰구간의 수렴을 위해 요구되는 총 표본의 크기도 증가하게 된다. 경험상, 신뢰할 수 있는 수준의 붓스트랩 신뢰구간을 유도하기 위해서는 각 셀 당 평균 10개 이상의 표본이 요구되었다. 우리는 여기서 BOOT1 신뢰구간과 BOOT2 신뢰구간의 특징을 살펴보고자 한다. 붓스트랩 추정량과 붓스트랩 분산을 계산한 후 단순히 정규분포의 분위수로 신뢰구간을 구축한 BOOT1 신뢰구간은 [표 3.1]과 [표 3.2]를 통하여 BOOT2 보다 더 나은 방법처럼 보이며, 표본의 크기가 작음에도 불구하고, 주어진 명목수준에 상당히 근접해 있음을 볼 수 있다. 그런데, 표본의 크기가 큰 총 표본이 320인 경우를 살펴보면 BOOT1 신뢰구간보다 BOOT2 신뢰구간이 신뢰구간의 평균 길이도 더 짧으며, 주어진 명목수준에 더 근사한 방법임을 알 수 있다. 이러한 경향은 사전분포가 정규분포일 때도 마찬가지로 확인할 수 있었다. 즉 주축통계량을 사용한 분포가 표본이 커짐에 따라 정규근사를 사용한 붓스트랩 방법보다 실제분포에 더 근사함을 알 수 있다.

4. 결론

본 연구에서는 평균포함확률 측면에서 공통오즈비의 신뢰구간 구축에 대해 살펴보았다. 특히, 기존의 신뢰구간 외에, 표본 재집계 방법인 잭나이프 신뢰구간과 붓스트랩 신뢰구간의 특징을 살펴보았다. 본 연구를 통해 공통오즈비의 추정에 있어서도 정확신뢰구간이 지니는 보수성을 확인할 수 있었다. 물론, 이산형 자료에서 표본의 크기가 작은 경우 정확한 추론은 매우 중요한 역할을 하게 된다. 그런데 표본의 크기가 적당히 크면 근사이론에 기초한 통계적 추론이 경우에 따라 더 정확하게 작용한다는 점이다. 본 연구를 통해 다음 몇 가지 결과와 제안을 제시하고자 한다.

첫째, 정확신뢰구간이나 $\log(\phi)$ 를 사용한 잭나이프 신뢰구간은 과대추정의 경향이 존재한다. 둘째, 정확신뢰구간의 대안으로 중간-p 조정 신뢰구간을 구축할 수 있다. 이 방법은 표본의 크기가 작은 경우에도 구축할 수 있으며, MH/RBG 방법에 비해 비교적 주어진 명목수준에 잘 근사하는 방법이다. 셋째, PETO 신뢰구간은 하향추정의 경향이 있으며, 추정치가 모수를 대부분 낮은 수준에서 추정하는 불일치 통계량이다. 넷째, 잭나이프 통계량은 Pigeot(1991)이 언급한 것처럼 기존의 다른 방법에 비해 표준오차가 작으나 ϕ 를 중심으로 신뢰구간을 구축했을 때 하향추정을, $\log(\phi)$ 를 중심으로 신뢰구간을 구축하면 심각한 상향추정의 문제를 발생시킨다. 또한 잭나이프 방법을 사용한 신뢰구간에서 근본적으로 한 셀의 값이 0일 때 신뢰구간 구축에 계산상의 어려움이 따를 수 있다. 다섯째, 주축통계량을 사용한 붓스트랩 방법은 표본의 크기가 작을 때는 하향추정의 경향이 있으나, 표본의 크기가 커짐에 따라 상당히 안정적으로 신뢰구간을 구축하여 주며, 정규근사에 의존하는 Mantel-Haenszel 통계량의 이론적 어려움을 피하게 해준다. 특히, 층의 수가 작으면서, 표본의 크기가 클 때, 평균포함확률이 주어진 명목수준에 상당히 빠른 속도로 일치한다는 점에서 모수에 대한 어떤 정보가 주어지지 않은 경우에 비모수적 방법의 대안으로 적절히 사용하길 권한다.

본 연구에 사용된 정확신뢰구간은 StatXact(1997)에서 계산할 수 있으며, 본 연구에서는 FORTRAN을 사용하여 프로그램 하였다.

참고문헌

- [1] 광민정, 박종태, 정형철, 김희규 (2002). 층화된 K 2×2 분할표에서 붓스트랩 방법을 이용한 공통오즈비의 추정과 비교. <평택대학교 논문집>, 제 16권, 367-377.
- [2] Agrest, A. and Coull, B.A. (1998) Approximate is better than exact for interval estimation of binomial proportions. *The American Statistician*, Vol. 52, 119-126.
- [3] Breslow, N.E. (1981). Odds ratio estimators when the data are sparse. *Biometrika*, Vol. 68, 73-84.
- [4] Breslow, N.E. and Liang, KY. (1982). The variance of the Mantel-Haenszel odds ratio. *Biometrics*, Vol. 38, 943-952.
- [5] Emerson, J.D. (1994) Combining estimates of the odds ratio: the state of the art. *Statistical Methods in Medical Research*, Vol. 3, 157-178.
- [6] Flanders, W.D. (1985) A new variance estimator for the Mantel-Haenszel odds ratio. *Biometrics*, Vol. 41, 637-642.
- [7] Hauck, W.W. (1979). The large-scale variance of the Mantel-Haenszel estimator of a common odds ratio. *Biometrics* Vol. 35, 817-819.
- [8] Hirji, K.F. and Vollset, S.E. (1993) Computation exact distribution for several 2×2 tables. *Applied statistics*, 270-274.
- [9] Kim, D.H. and Jeong, H.C. (2000a) A simulation study for the confidence interval of p by using average coverage probability. *The Korean Communications in Statistics*, Vol. 7, No. 3, 859-869.
- [10] Kim, D.H. and Jeong, H.C. (2000b) A comparison of some approximate confidence interval for the poisson parameter. *The Korean Communications in Statistics*. Vol. 7, No. 3, 899-911.
- [11] Mehta, C.R., Patel, N.R. and Gray, R. (1985) Computing an Exact Confidence Interval for the Common Odds Ratio in Several 2 × 2 Contingency Tables. *Journal of the American Statistical Association*, Vol. 80, 969-973.
- [12] Mehta, C.R. and Walsh, S.J. (1992) Comparison of exact, mid-P, and Mantel-Haenszel confidence intervals for the common odds ratio across several 2×2 contingency tables. *The American Statistician*, Vol. 46, 146-150.
- [13] Pigeot, I. (1991) A Jackknife Estimator of a Combined Odds Ratio. *Biometrics*, Vol. 47, 373-381.
- [14] Phillips, A. and Holland, P.W. (1987) Estimators of the variance of the Mantel-Haenszel log-odds-ratio estimate. *Biometrics*, Vol. 43, 425-431.
- [15] Robins, J., Breslow, N.E. and Greenland, S. (1986) Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics*, Vol. 42, 311-323.
- [16] Sato, T. (1990) Confidence limits for the common odds ratio based on the asymptotic distribution of the Mantel-Haenszel estimator. *Biometrics* Vol. 46, 71-80.

- [17] Singh, K. (1981) On the asymptotic accuracy of Efron's bootstrap, *The Annals of Statistics*, Vol. 9, 1187-1195.
- [18] StatXact(1997) *A statistical package for exact nonparametric inference*, Cytel software corporation, 137 Eric street, Cambridge, MA.
- [19] Vollset, S.E., Hirji, K.F. and Elashoff R.M. (1991) Fast computation of exact confidence limits for the common odds ratio in a series of 2×2 contingency tables. *Journal of the American Statistical Association*, Vol. 86, 404-409.
- [20] Woodroffe, M. and Jhun, M. (1989) Singh's theorem in the lattice case. *Statistics Probability Letters*, Vol. 7, 201-205.

[2003년 7월 접수, 2003년 12월 채택]