

# Confidence Interval for Sensitive Binomial Attribute : Direct Question Method and Indirect Question Method

Jea-Bok Ryu<sup>a,1</sup>

<sup>a</sup>Department of Statistics, College of Science & Engineering, Cheongju University

(Received November 18, 2014; Revised January 23, 2015; Accepted February 2, 2015)

---

## Abstract

We discuss confidence intervals for sensitive binomial attributes obtained by a direct question method and indirect question method. The Randomized Response Technique(RRT) by Warner (1965) is an indirect question method that uses a randomization device to reduce the response burden of respondents. We used the mean coverage probability (MCP), root mean squared error (RMSE), and mean expected width (MEW) to compare the confidence intervals by the two methods. The numerical comparisons indicated found that the MEW of RRT is too large and the RRT is so conservative that the MCP exceeds a nominal level( $\alpha$ ); therefore, it is necessary to complement these problem in order to increase the utility of the indirect question method.

Keywords: Sensitive binomial attribute, confidence interval, direct question method, indirect question method.

---

## 1. 서론

이항비율에 대한 신뢰구간 추정의 문제는 오랜 동안 통계분야에서 중요하게 다루어져 왔다. 일반적으로 정규근사이론을 이용한 Wald 방법이 사용되고 있으나 이의 적절성에 대한 문제가 제기되면서 Wald 신뢰구간을 대체할 많은 방법들이 연구되었다 (Vollset, 1993; Newcombe, 1998; Agresti와 Coull, 1998; Brown 등, 2001, 2002; Ryu와 Lee, 2006; Ryu, 2009). 특히 이항모집단의 비율  $p$ 가 작은 경우에는 극단값이 신뢰구간의 포함확률이나 신뢰구간의 길이에 영향을 주기 때문에 극단값을 조정하는 방법들이 제시되었다 (Agresti와 Coull, 1998; Brown 등, 2001; Ryu, 2010, 2011).

이항비율  $p$ 가 작은 경우는 민감한 사항에 대한 조사에서 많이 발생한다. 예를 들면, 약물남용, 탈세 여부, 혼전 성경험, 희귀질병 감염여부 등과 같이 응답자들이 응답하기 꺼려하는 조사의 경우다. 이러한 조사에서는 응답자들이 부담을 느껴서 응답을 거부하거나 거짓 응답을 하는 경향이 있다. 특히 개인의 신분보장을 매우 중요하게 생각하는 최근에는 민감한 사항에 대한 직접질문방식이 현실적으로 어렵다. 이러한 현실적 문제를 고려해서 Warner (1965)는 간접질문방식인 확률화응답기법(Randomized Response Technique; RRT)을 제안하였다. 확률화응답기법은 회전판, 주사위, 카드 등과 같은 확률

---

This work was supported by the research grant of Cheongju University in 2014.

<sup>1</sup>Department of Statistics, College of Science & Engineering, Cheongju University, 298 Daeseong-ro, Cheongwon-gu, Cheongju, Chungbuk 363-764, Republic of Korea. E-mail: jbryu@cju.ac.kr

장치를 이용하여 선택된 질문에 응답하게 하는 방식이다. 어떤 질문이 선택되었는지를 조사자는 알 수 없고 응답자만이 알 수 있어 응답자의 신분이 보장된다. 물론 (민감한) 질문이 선택될 확률  $T$ 는 사전에 조정해 줄 수 있다. Warner의 RRT는 민감한 질문과 관련된 두 질문을 확률장치를 이용해서 민감한 이항특성에 대한 추정 문제를 다루고 있다. Warner는 간접질문방식을 제안하면서 민감한 이항특성에 관한 추정치와 추정치의 분산을 유도하였다. 그러나 이들은 확률장치에서 민감한 질문이 선택될 확률  $T$ 에 의존하게 된다. 또한 그는 직접질문법에 대한 간접질문법의 효율을 평균제곱오차로 비교하였다. 한편 Greenberg 등 (1969)은 민감한 질문과 관련이 없는 질문을 사용하는 무관질문모형을 제시하였다. Warner 이후 민감한 사항에 대한 간접질문법에 대한 연구는 수 없이 많다. 확률화응답기법들에 대한 전반적인 내용들은 Ryu 등 (1993), Chaudhuri (2011) 등에 정리되어 있다.

민감한 이항특성에 대한 추정이 직접질문법이든 간접질문법이든지 간에 이용자들은 각 방법에 의한 추정치를 이용한 신뢰구간의 문제에 관심을 가지게 된다. 응답자들이 응답을 꺼리는 민감한 특성은 통상적으로 출현 비율이 낮으므로 결국에는 낮은 이항비율에 대한 신뢰구간의 문제가 된다. 따라서 본 연구에서는 민감한 이항특성에 대한 직접질문법과 간접질문법에 의한 신뢰구간 추정 문제를 다루고자 한다. 두 방법에 의한 추정이 매우 다양하므로, 여기서는 대표적인 방법들을 비교해 본다. 즉, 민감한 이항특성에 대해 Wald의 직접질문법과 Warner (1965)의 간접질문법을 이용해서 구한 신뢰구간을 비교한다. 두 신뢰구간을 비교하기 위해서 신뢰구간의 평가기준으로 널리 사용되고 있는 평균포함확률(mean coverage probability; MCP), 평균제곱오차의 제곱근(root mean squared error; RMSE), 그리고 평균기대폭(mean expected width; MEW)을 사용한다.

논문의 구성은 직접질문법과 간접질문법의 대표적인 Wald 방법과 Warner 방법을 각각 2장과 3장에서 살펴본다. 4장에서는 이항특성(비율)에 대한 신뢰구간 추정의 타당성을 평가하는 기본적인 3가지 기준인 평균포함확률, 평균제곱오차의 제곱근, 그리고 평균기대폭들을 다룬다. 그리고 5장에서는 간접질문법과 직접질문법에 의한 신뢰구간에 대해 3가지 기준을 사용하여 수치적 비교를 한다.

## 2. 직접질문법

민감한 특성에 대한 직접질문(direct question; DR)법은 “당신은 민감한 속성  $A$ 를 가지고 있습니까?”라는 질문을 응답자들에게 하고 이에 대해 응답자들이 직접 응답하는 방식이다. 따라서 이항모집단의 민감한 비율  $p$ 에 대한 최우추정치(MLE)를 구하고, 점근적 정규분포이론에 근거해 신뢰구간을 구한다.

이항비율에 대해 직접질문을 이용해서 신뢰구간을 추정하는 방법은 매우 다양하다. 일반적으로 널리 사용되고 있는 방법은 Wald 방법이지만, 이 방법의 부적합성에 대한 상세한 의견과 대체 방법들은 Ryu (2009)를 참고하면 된다. 하지만 본 연구에서는 직접질문법과 간접질문법의 대표적인 방법에 의한 신뢰구간을 비교하는 것이므로 직접질문에서는 Wald 방법을 사용한다.

Wald에 의한 이항비율  $p$ 에 대한 신뢰구간의 중심은 모비율에 대한 최우추정치이고 상한과 하한의 계산에 추정치의 추정표준오차가 사용된다. 즉,

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad (2.1)$$

여기서  $z_{\alpha/2}$ 는 표준정규분포의  $1 - \alpha/2$  분위수이다.

## 3. 간접질문법

민감한 질문에 대해 응답자들이 응답부담을 갖게 되어 응답을 회피하거나 거짓 응답하는 경향이 있다.

따라서 직접 질문하는 대신에 확률장치를 이용해서 응답하는 간접질문법이 Warner에 의해 처음으로 제시되었다. Warner의 간접질문법을 확률화회응답기법(randomized response technique; RRT)이라 부르게 되었고, 이에 대한 다양한 간접질문법들이 제시되었다. Warner의 RRT는 민감한 속성을 갖는 모집단 비율을 추정하기 위하여 다음과 같은 민감한 속성과 관련이 있는 두 개의 질문을 사용한다.

질문1 : “당신은 민감한 속성  $A$ 를 가지고 있습니까?”

질문2 : “당신은 민감한 속성  $A$ 를 가지고 있지 않습니까?”

단순임의복원추출된  $n$ 명의 응답자들은 확률장치에 의해서 선택된 질문에 “예” 또는 “아니오”로만 응답한다. 질문1이 선택될 확률을  $T$ , 질문2가 선택될 확률을  $1 - T$ 라 하면 응답자가 “예”라고 응답할 확률은 다음과 같다.

$$k = Tp + (1 - T)(1 - p) \quad (3.1)$$

여기서,  $p$ 는 민감한 속성을 갖는 모집단 비율이다. 만약  $T = 1$ 이면 직접질문의 경우가 된다.  $n$ 명의 응답자 중에서 “예”라고 응답한 사람의 수가  $n'$ 이면, 민감한 속성  $A$ 를 갖는 모집단비율  $p$ 의 추정량  $\hat{p}_w$ 와 분산  $V(\hat{p}_w)$ 는 다음과 같다.

$$\hat{p}_w = \frac{\hat{k} - (1 - T)}{2T - 1}, \quad \hat{k} = \frac{n'}{n}, \quad T \neq \frac{1}{2}, \quad (3.2)$$

$$V(\hat{p}_w) = \frac{p(1 - p)}{n} + \frac{1}{n} \left[ \frac{1}{16(T - 0.5)^2} - \frac{1}{4} \right]. \quad (3.3)$$

식 (3.2)와 식 (3.3)으로부터 정규근사에 의한 신뢰구간을 구하면 다음과 같다.

$$\hat{p}_w \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_w(1 - \hat{p}_w)}{n} + \frac{1}{n} \left[ \frac{1}{16(T - 0.5)^2} - \frac{1}{4} \right]}. \quad (3.4)$$

#### 4. 신뢰구간의 평가기준

이항비율에 대한 신뢰구간을 구하는 문제는 많은 연구자들에 의해서 이루어지고 있다. 본 연구에서는 민감한 특성을 갖는 이항비율을 추정하는 데 직접질문법(DR)과 확률장치를 이용한 간접질문법(RRT)을 사용하였을 경우의 신뢰구간을 비교하고자 한다. 어떠한 방법에 의하든 간에 신뢰구간의 적절성을 평가하는 것이 필요하다. 일반적인 평가기준으로 포함확률과 포함확률을 모든 모수공간에 대해 평균한 평균포함확률이 널리 사용된다. 그리고 포함확률이 명목수준으로부터 얼마나 떨어져 있는가를 나타내는 평균제곱오차의 제공근도 중요한 척도이다. 또한 포함확률을 보완해주기 위해서 신뢰구간의 상한과 하한의 차이인 기대폭을 모든 모수공간에 대해 평균한 평균기대폭도 사용된다.

##### 4.1. 평균포함확률

포함확률(coverage probability)은 특정 모수가 신뢰구간에 포함될 확률로 다음과 같이 정의된다.

$$C_n(p) = \sum_{x=0}^n I(x, p) \binom{n}{x} p^x (1 - p)^{n-x}, \quad (4.1)$$

여기서  $I(x, p)$ 는  $X = x$ 일 때 신뢰구간이  $p$ 를 포함하고 있으면 1이고 그렇지 않으면 0이다. 포함확률은 이항모수  $n$ 과  $p$ 의 함수로 이들의 값에 따라 변한다. 그런데 이항분포는 이산형이기 때문에  $n$ 과  $p$ 의

조그만 변화에도 포함확률의 변동이 매우 크다. 그런데, 모수값이 어느 위치에 있는지를 모르므로 모든 모수공간에 대해서 포함확률을 평균한 평균포함확률(mean coverage probability; MCP)을 이용하는 것이 바람직하다. 즉,

$$\bar{C}_n(p) = \int_0^1 C_n(p) f(p) dp, \quad (4.2)$$

여기서  $f(p)$ 는 모비율  $p$ 에 대한 밀도함수로 균등분포를 사용한다.

#### 4.2. 평균제곱오차의 제공근

한편 포함확률이 명목수준( $\alpha$ )으로부터 얼마나 떨어져 있는가를 나타내는 척도로 다음과 같은 평균제곱오차의 제공근(root mean square error; RMSE)을 사용한다.

$$\text{RMSE} = \sqrt{\int_0^1 (C_n(p) - \alpha)^2 dp}. \quad (4.3)$$

#### 4.3. 평균기대폭

포함확률은 신뢰구간의 상한과 하한의 폭에 의존하는데, 폭이 넓으면 포함확률이 커지고 반대로 폭이 좁으면 작아지는 경향이 있다. 신뢰구간의 폭이 너무 넓으면 포함확률은 커지지만 추정의 의미는 없게 된다. 따라서 포함확률이 명목수준에 근사하면서도 신뢰구간의 폭이 가능한 한 작은 것이 바람직하다. 신뢰구간의 상한과 하한의 차에 의한 기대폭(expected width; EW)은 다음과 같이 구한다.

$$E_n(p) = \sum_{x=0}^n (U(x) - L(x)) \binom{n}{x} p^x (1-p)^{n-x}, \quad (4.4)$$

여기서  $U(x)$ 와  $L(x)$ 는 각각 신뢰구간의 상한과 하한이다. 신뢰구간의 기대폭을 모든 모수공간에 대해 평균한 평균기대폭(mean expected width; MEW)은 다음과 같다.

$$\bar{E}_n(p) = \int_0^1 E_n(p) dp. \quad (4.5)$$

### 5. 수치적 비교

직접질문법과 간접질문법을 이용하여 구한 민감한 이항비율에 대한 신뢰구간을 비교하기 위해서 4장에서 다룬 MCP, RMSE, MEW를 사용한다.

직접질문법과 간접질문법에 의한 신뢰구간은 각각 식 (2.1)과 식 (3.4)를 이용하여 구한다. Warner의 간접질문법에 의한 신뢰구간은 확률장치를 이용하게 되므로 민감한 질문이 선택될 확률  $T$ 에 따라 달라진다. 따라서 직접질문법과 간접질문법에 의한 신뢰구간을 비교하기 위해서 간접질문법의 경우  $T = 0.6, 0.7, 0.8, 0.9$ 인 경우를 다룬다.  $T = 1$ 인 경우는 직접질문법이 된다.

Table 5.1, Table 5.2, 그리고 Table 5.3은 직접질문법과 간접질문법을 사용한 경우에 계산된 MCP, RMSE, MEW를 나타낸다.

신뢰구간을 평가하는 기준에서 가장 중요한 요소인 평균포함확률(MCP)은 직접질문법의 경우 표본 수가 증가함에 따라 MCP도 증가하나, 간접질문법인 Warner의 RRT는 표본수가 증가해도 그리고 민감한 질문이 선택될 확률  $T$ 의 변화에도 차이가 거의 없이 1에 가깝다. 통상적인 RRT에 대한 연구에서

**Table 5.1.** MCP for the 95% confidence intervals by direct question method(DR) and indirect question method(RRT) with  $n = 10, 30, 100, 500$  and different  $T$ .

$n$	DR( $T = 1$ )	RRT			
		$T$			
		0.6	0.7	0.8	0.9
10	0.7692893	0.9999	0.9998998	0.9987937	0.9788233
30	0.8749269	0.9999	0.9998997	0.9992623	0.9877289
100	0.9222814	0.9999	0.9998997	0.9994098	0.9905677
500	0.9430100	0.9999	0.9998998	0.9994570	0.9914889

**Table 5.2.** RMSE for the 95% confidence intervals by direct question method(DR) and indirect question method(RRT) with  $n = 10, 30, 100, 500$  and different  $T$ .

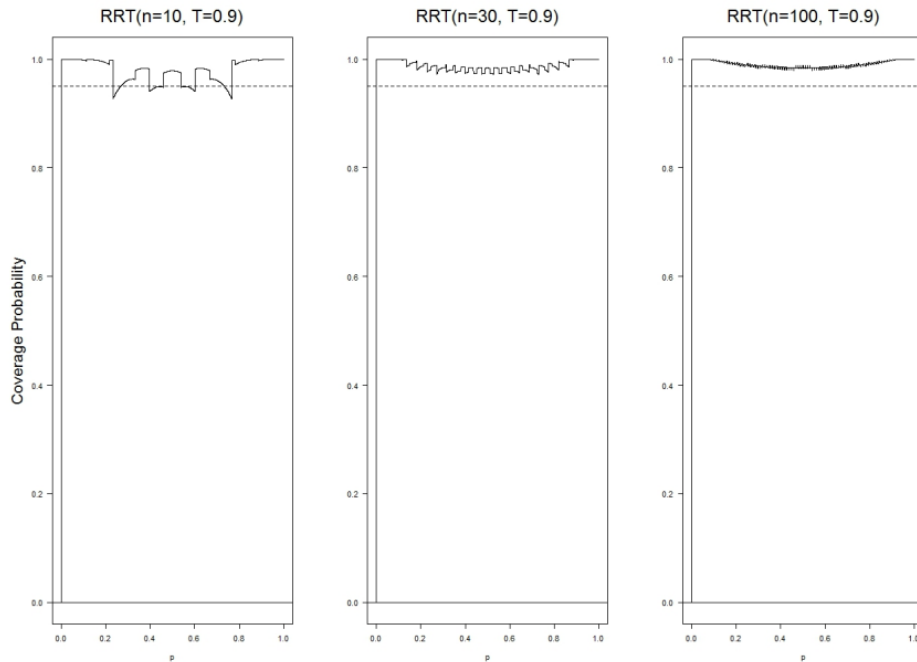
$n$	DR( $T = 1$ )	RRT			
		$T$			
		0.6	0.7	0.8	0.9
10	0.28631680	0.05089195	0.05089171	0.04983497	0.03729940
30	0.16832010	0.05089195	0.05089165	0.05027056	0.03997960
100	0.09312201	0.05089195	0.05089169	0.05041268	0.04219189
500	0.04189856	0.05089195	0.05089171	0.05045855	0.04299216

**Table 5.3.** MEW for the 95% confidence intervals by direct question method(DR) and indirect question method(RRT) with  $n = 10, 30, 100, 500$  and different  $T$ .

$n$	DR( $T = 1$ )	RRT			
		$T$			
		0.6	0.7	0.8	0.9
10	0.4034284	0.9999000	0.9083038	0.7021720	0.54340690
30	0.2662789	0.9818065	0.6731368	0.4788381	0.35781270
100	0.1518312	0.7339904	0.4203623	0.2847022	0.20714210
500	0.0686743	0.3881257	0.2024621	0.1329723	0.09502102

RRT 방법들의 효율이  $T$ 의 변화에 민감한 것과는 대조적이다. 즉, 간접질문을 사용한 신뢰구간 추정에서 MCP는 민감한 질문이 선택될 확률  $T$ 에 별 영향을 받지 않는다는 것이다. 반면에 평균제곱오차의 제공근인 RMSE는 직접질문법의 경우 표본수가 증가함에 따라 감소하나 간접질문법의 경우는 표본수가 증가해도 변화가 거의 없지만  $T$ 가 커질수록 감소한다. 이는 직접질문법의 경우 포함확률은  $n$ 이 작을 때  $p$ 의 변화에 따라 변동이 매우 심하지만, 간접질문법의 경우는 그렇지 않기 때문이다. 직접질문법에 비해 간접질문법의 RMSE는 전반적으로 매우 작다. 한편 MEW는 포함확률과 연관이 되므로 직접질문법보다 간접질문법의 경우가 전체적으로 크고, 두 방법 모두 표본이 증가하면 감소한다. 그리고 간접질문법의 경우는 민감한 질문이 선택될 확률  $T$ 가 증가하면 감소한다.

Table 5.1에 의하면 간접질문법의 MCP는 표본수나 민감한 질문이 선택될 확률  $T$ 에 관계없이 매우 높다. Figure 5.1은  $n = 10, 30, 100$ 이고  $T = 0.9$ 인 경우 이항비율  $p$ 의 변화에 따른 포함확률(CP)을 보여주고 있는데, 전반적으로 포함확률이 명목수준( $\alpha$ )을 크게 초과하고 있음을 알 수 있다. 모든  $p$ 의 값에 대해서 포함확률이 명목수준을 훨씬 초과하고 있어서 간접질문법에 의한 신뢰구간 추정은 매우 보수적이다.  $T$ 가 작을 때도 MCP와 포함확률이 모두 1에 근접하기 때문에 Figure 5.1에서는  $T = 0.9$ 인 경우만을 보여준다. 이항비율에 대한 신뢰구간 추정에서 바람직한 것은 평균포함확률이 명목수준에 근사하고 포함확률도  $p$ 의 모든 구간에 대해 명목수준에 근사한 것이다 (Newcombe, 1998; Brown 등, 2001;



**Figure 5.1.** Coverage Probability for the 95% confidence intervals by indirect question method(RRT) with  $n = 10, 30, 100$  and  $T = 0.9$ .

Ryu, 2009). 따라서 간접질문법인 RRT를 보건, 의학, 공학 등 이항비율에 대한 신뢰구간을 이용하는 분야에 적용하기 위해서는 직접질문법에서 최근에 많이 연구되고 있는 평균포함확률을 명목수준에 근사하도록 조정해 주는 방법에 대한 연구가 필요하다.

Land 등 (2012)는 민감한 속성을 갖는 사람들의 수가 적기 때문에 많은 수의 표본이 요구되지만, 통신 체계의 빠른 발전으로 전화조사나 인터넷조사 등으로 많은 표본에 대한 간접질문법인 RRT가 가능하다고 주장하였다. 또한 RRT에 대한 불신이나 이해의 어려움도 교육과 홍보로 향후 사용이 충분하다고 언급하고 있다. 그러나 본 연구의 수치적 비교에 의하면 표본 수가 증가해도 간접질문법인 RRT에 의한 추정치의 신뢰구간에 대한 MCP나 MEW는 만족스러운 결과를 얻을 수 없다. 하지만 민감한 질문이 선택될 확률  $T$ 를 낮게 하더라도 MCP나 MEW가 별 영향을 받지 않으므로 민감한 질문에 대한 조사에 응답자들의 신분을 보장해서 보다 정확하고 신뢰할 수 있는 응답을 얻을 수 있다.

## 6. 결론 및 토의

민감한 사항에 대한 조사에서 직접질문을 하면 응답자들이 응답 부담을 느껴 거짓 응답을 하거나 응답을 기피하는 현상이 발생한다. 최근에는 민감한 사항들에 대한 조사가 빈번해지지만, 사회적으로 개인의 비밀을 중시하는 경향에 따라 직접질문법의 사용은 점점 어려워지고 있다. 이를 보완하기 위해서 확률장치를 이용한 간접질문방식인 확률화응답기법(RRT)이 1965년 Warner에 의해 처음 제시되었다.

직접질문이나 간접질문을 이용하여 민감한 이항비율에 대한 추정과 이를 바탕으로 점근적 정규이론에 근거한 신뢰구간 추정은 실제적으로 많은 분야에서 활용되고 있다.

본 연구에서는 민감한 이항특성에 대한 신뢰구간을 직접질문과 간접질문을 사용하여 구했을 경우, 이

들 신뢰구간의 적절성을 수치적 비교를 통해서 살펴보았다. 현실적으로는 직접질문법보다는 간접질문법이 응답자들의 비밀을 보장해주어 솔직한 응답이 가능하고 무응답을 줄여서 보다 정확한 추정이 가능하다. 그러나 두 방법에 의한 신뢰구간의 적절성을 평가하는 평균포함확률(MCP), 평균제곱오차의 제곱근(RMSE), 그리고 평균기대폭(MEW)을 비교한 결과, 표본수가 커질 때 직접질문법의 MCP는 증가하나 간접질문법인 RRT는 표본수가 증가하거나 민감한 질문이 선택될 확률  $T$ 의 변화에도 차이가 거의 없이 1에 가깝다. 또한 RMSE는 직접질문법의 경우 표본수가 증가하면 감소하나 간접질문법에서는 표본수가 증가해도 변화가 거의 없다. 다만,  $T$ 가 커질수록 간접질문법의 RMSE 감소한다. 전반적으로 직접질문법에 비해 간접질문법의 RMSE는 매우 작다. 한편 MEW는 직접질문법과 간접질문법의 경우 모두 표본이 증가하면 감소하고 간접질문법의 경우는 민감한 질문이 선택될 확률  $T$ 가 증가하면 감소한다. 그러나 전체적으로 직접질문법보다 간접질문법의 경우 MEW가 훨씬 크다.

간접질문법인 RRT가 비록 응답자들의 신분을 보장하여, 응답자들이 솔직하게 응답할 수 있게 하고, 무응답률을 낮출 수 있는 바람직한 방법으로 50여 년 동안 많은 연구들이 이루어지고 있다. 하지만 본 연구에 의하면 RRT를 이용한 신뢰구간 추정에 한계가 있다는 것을 알 수 있다. 하나의 특이점은 RRT에 대한 기존의 연구와는 달리 이항비율에 대한 신뢰구간 추정에서 민감한 질문이 선택될 확률  $T$ 가 MCP에 별 영향을 미치지 않으므로 응답자의 신분 보장에 매우 유용하다. 반면에 RRT를 이용한 신뢰구간의 MCP와 명목수준과의 차이는 더 심하고 포함확률도 1에 근접하게 되어 매우 보수적이며, MEW도 직접질문에 비해 상당히 크므로 이들을 보완해 주는 연구가 필요하다.

## References

- Agresti, A. and Coull, B. A. (1998). Approximate is better than “Exact” for interval estimation of Binomial proportions, *The American Statistician*, **52**, 119–126.
- Brown, L. D., Cai, T. T. and DasGupta, A. (2001). Interval estimation for a binomial proportion (with discussion), *Statistical Science*, **16**, 101–133.
- Brown, L. D., Cai, T. T. and DasGupta, A. (2002). Confidence intervals for a binomial proportion and asymptotic expansions, *The Annals of Statistics*, **30**, 160–201.
- Chaudhuri, A. (2011). *Randomized Response and Indirect Questioning Techniques in Surveys*, Chapman & Hall/CRC, New York.
- Greenberg, B. G., Abul-Ela, Abdel-Latif A., Simmons, W. R. and Horvitz, D. G. (1969). The unrelated question randomized response model; Theoretical framework, *Journal of the American Statistical Association*, **64**, 520–539.
- Land, M., Singh, S. and Sedory, S. (2012). Estimation of a rare sensitive attribute using Poisson distribution, *Statistics*, **46**, 351–360.
- Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: Comparison of seven methods, *Statistics in Medicine*, **17**, 857–872.
- Ryu, J. B. (2009). A short consideration of binomial confidence interval, *Communications of the Korean Statistical Society*, **16**, 731–743.
- Ryu, J. B. (2010). The effect of adjusting the extreme values in Wald confidence interval, *Journal of Research Institute of Industrial Sciences*, **28**, 29–34.
- Ryu, J. B. (2011). The influence of extreme value in binomial confidence interval, *Communications of the Korean Statistical Society*, **18**, 615–623.
- Ryu, J. B., Hong, K. H. and Lee, G. S. (1993). *Randomized Response Model*, Freedom Academy, Seoul.
- Ryu, J. B. and Lee, S. J. (2006). Confidence intervals for a low binomial proportion, *The Korean Journal of Applied Statistics*, **19**, 217–230.
- Vollset, S. E. (1993). Confidence intervals for a binomial proportion, *Statistics in Medicine*, **12**, 809–824.
- Warner, S. L. (1965). Randomized response: A survey technique for elimination evasive answer bias, *Journal of the American Statistical Association*, **60**, 63–69.

# 민감한 이항특성에 대한 신뢰구간 : 직접질문법과 간접질문법

류제복<sup>a,1</sup>

<sup>a</sup>청주대학교 이공대학 통계학과

(2014년 11월 18일 접수, 2015년 1월 23일 수정, 2015년 2월 2일 채택)

---

## 요약

본 연구에서는 민감한 이항특성에 대한 신뢰구간 추정에 직접질문법과 간접질문법을 사용한다. 간접질문법으로 응답자들의 응답부담을 줄여주기 위해서 확률장치를 사용하는 Warner (1965)의 확률화응답기법(RRT)을 고려한다. 두 방법에 의한 신뢰구간을 비교하기 위해서 평가기준으로 평균포함확률(MCP), 평균제곱오차의 제곱근(RMSE), 그리고 평균기대폭(MEW)을 사용한다. 수치적 비교 결과 RRT의 MCP가 명목수준( $\alpha$ )을 크게 초과하여 보수적이고 MEW도 매우 크다. 따라서 이들을 보완해 주어야 실제적으로 간접질문법의 유용성을 높일 수 있다.

주요용어: 민감한 이항특성, 신뢰구간, 직접질문법, 간접질문법.

---

이 논문은 2014년도에 청주대학교 산업과학연구소가 지원한 학술연구조성비(특별연구과제)에 의해 연구되었음.

<sup>1</sup>(363-764) 충북 청주시 청원구 대성로 298, 청주대학교 이공대학 통계학과. E-mail: jbryu@cju.ac.kr