Journal of Korean
Data & Information Science Society
2002, Vol. 13, No.1 pp. 25 34

A Study on the Multiple Trial of Unrelated Question Models[1)]

Gi Sung Lee[2)] · Ki Hak Hong[3)]

Abstract

In this paper, we proposed multiple trial unrelated question models that more efficient by reducing the variance of the estimate than single trial unrelated question models investigated by Greenberg et al.'s(1969) and Kim et al.'s(1992) and Lee & Hong's(1998).

Keywords : 개선된 무관질문모형, 2단계 무관질문모형, 중복시행모형

1. 서론

사회 여러 분야의 조사에서 응답자들이 민감하거나 개인적인 이해와 관계되는 질문을 직접 받았을 경우 편향이 발생하게 되는데, 이러한 편향을 줄이기 위하여 Warner (1965)는 확률장치를 이용하여 간접응답을 얻는 확률화응답모형(randomized response model ; RRM)을 처음으로 제시하였다. 그 후 응답자의 신분이나 비밀을 보장함으로서 응답자로부터 민감한 질문에 보다 더 정확한 정보를 얻을 수 있는 확률화응답모형은 많은 학자들에 의해 연구 발전되었으며, 특히 Greenberg et al.(1969)은 무관질문모형(unrelated question model)을 제안하였다. 최근에 Mangat-Singh(1990)은 2단계 관련질문모형을 제안하였으며, 김종호외 2인(1992)은 이 모형을 2단계 무관질문모형으로 발전시켰다. 그리고, Mangat(1994)은 Mangat-Singh의 2단계 관련질문모형에서 사용한 2개의 확률장치를 1개로 줄여 그 사용 절차를 좀 더 단순화 한 개선된 관련질문모형을 제안하였다. 또한, 이기성과 홍기학(1998)은 Mangat의 개선된 관련질문모형을 무관질문모형으로 발전시킨 개선된 무관질문모형을 제안하였다. 그러나, 이러한 확률화응답모형들은 조사자에게 응답자의 비밀을 노출시키지 않는다는 장점을 가지고 있으나 직접질문에 의한 조사에 비하여 추정량의 분산이 상대적으로 커진다는 단점을

---

1. This paper was supported by Woosuk University
2. Associate Professor, Division of Computer and Information Science, Woosuk University, 490 Hujung-ri, Samrye-up, Wanju-gun, Jeonbuk, 565-701, Korea
E-mail : gisung@woosuk.ac.kr
3. Associate Professor, Department of Computer Science, Dongshin University, 252 Daeho- dong, Naju, Chonnam, 520-714, Korea

가지고 있다. 이러한 점을 보완하기 위하여 Liu-Chow(1976)는 Warner의 모형을 중복하여 사용함으로서 표본의 크기를 증가시키지 않고서도 추정량의 분산을 줄일 수 있는 중복시행모형(multiple trial model)을 제안하였다.

본 논문에서는 Greenberg et al.의 무관질문모형과 김종호외 2인의 2단계 무관질문모형, 그리고 이기성과 홍기학의 개선된 무관질문모형을 중복시행모형으로 확장하여 Liu-Chow의 중복시행모형과 그 효율성을 비교하고자 한다.

2. 무관질문모형들의 단일시행

이 장에서는 무관한 속성 Y에 대한 모비율 π_y를 알고 있을 때, 민감한 속성에 대한 모비율 π를 추정하는 무관질문모형들 즉, Greenberg et al.의 무관질문모형과 2단계 무관질문모형, 그리고 개선된 무관질문모형에 대하여 살펴보고자 한다.

2.1 Greenberg et al.의 무관질문모형

단순임의복원추출된 응답자들은 확률장치를 이용해서 다음과 같은 2개의 설문 중 선택된 질문에 대하여 응답한다.

설문 1 : 당신은 민감한 속성 A를 가지고 있습니까?
설문 2 : 당신은 무관한 속성 Y를 가지고 있습니까?

여기서, 설문 1이 선택될 확률은 p이고, 설문 2가 선택될 확률은 $1-p$이다. 이 때, 응답자들은 확률장치에 의해서 선택된 설문에 대해 "예" 또는 "아니오"라고 응답한다. 응답자에 의해 이루어지는 이러한 모든 절차를 조사자는 관찰할 수 없다.

단순임의복원추출된 n명의 응답자들 중에서 "예"라고 응답한 사람의 수를 n'이라 하면, 민감한 속성에 대한 모비율 π의 최대우도추정량 $\hat{\pi}_{u_1}$는 다음과 같다.

$$\hat{\pi}_{u_1} = \frac{\frac{n'}{n} - (1-p)\pi_y}{p} . \tag{2.1}$$

한편, 최대우도추정량 $\hat{\pi}_{u_1}$는 π의 비편향추정량이며, 그 분산은 다음과 같다.

$$V(\hat{\pi}_{u_1}) = \frac{\pi(1-\pi)}{n} + \frac{(1-p)[p(1-2\pi_y)\pi + \pi_y\{1-(1-p)\pi_y\}]}{np^2} . \tag{2.2}$$

2.2 2단계 무관질문모형

단순임의복원추출된 응답자들은 1단계로 확률장치 R_1을 이용해서 다음과 같은 2개

의 설문 중 선택된 질문에 대하여 응답한다.

설문 1 : 당신은 민감한 속성 A 를 가지고 있습니까?
설문 2 : 확률장치 R_2로 가시오.

여기서, 설문 1이 선택될 확률은 T이고, 설문 2가 선택될 확률은 $1-T$이다.

한편, 설문 2가 선택된 경우 2단계로 확률장치 R_2를 이용해서 다음과 같은 2개의 설문 중 선택된 질문에 대하여 응답한다.

설문 1 : 당신은 민감한 속성 A 를 가지고 있습니까?
설문 2 : 당신은 무관한 속성 Y를 가지고 있습니까?

여기서, 설문 1이 선택될 확률은 p이고, 설문 2가 선택될 확률은 $1-p$이다. 이 때, 응답자들은 확률장치에 의해서 선택된 설문에 대해 “예” 또는 “아니오”라고 응답한다.

단순임의복원추출된 n명의 응답자들 중에서 “예”라고 응답한 사람의 수를 n'이라 하면, 민감한 속성에 대한 모비율 π의 최대우도추정량 $\hat{\pi}_{u_2}$는 다음과 같다.

$$\hat{\pi}_{u_2} = \frac{\frac{n'}{n} - (1-T)(1-p)\pi_y}{T + (1-T)p} \quad . \tag{2.3}$$

한편, 최대우도추정량 $\hat{\pi}_{u_2}$는 π의 비편향추정량이며, 그 분산은 다음과 같다.

$$V(\hat{\pi}_{u_2}) = \frac{\pi(1-\pi)}{n} + \frac{\{T+(1-T)p\}(1-T)(1-p)\{\pi(1-2\pi_y)+\pi_y\} + (1-T)^2(1-p)^2(1-\pi_y)\pi_y}{n\{T+(1-T)p\}^2} \quad . \tag{2.4}$$

2.3 개선된 무관질문모형

단순임의복원추출된 응답자들은 민감한 속성 A 를 가지고 있으면 “예”라고 응답하게 되고, 민감한 속성 A 를 가지고 있지 않으면 확률장치를 이용해서 다음과 같은 2개의 설문 중 선택된 질문에 대하여 응답한다.

설문 1 : 당신은 민감한 속성 A 를 가지고 있습니까?
설문 2 : 당신은 무관한 속성 Y를 가지고 있습니까?

여기서, 설문 1이 선택될 확률은 p이고, 설문 2가 선택될 확률은 $1-p$이다. 이 때, 응답자들은 확률장치에 의해서 선택된 설문에 대해 “예” 또는 “아니오”라고 응답한다.

단순임의복원추출된 n명의 응답자들 중에서 “예”라고 응답한 사람의 수를 n'이라 하면, 민감한 속성에 대한 모비율 π의 최대우도추정량 $\hat{\pi}_{u_3}$는 다음과 같다.

$$\hat{\pi}_{u_3} = \frac{\frac{n'}{n} - (1-p)\pi_y}{1-(1-p)\pi_y} . \tag{2.5}$$

한편, 최대우도추정량 $\hat{\pi}_{u_3}$는 π의 비편향추정량이며, 그 분산은 다음과 같다.

$$V(\hat{\pi}_{u_3}) = \frac{\pi(1-\pi)}{n} + \frac{(1-\pi)(1-p)\pi_y}{n\{1-(1-p)\pi_y\}} . \tag{2.6}$$

3. 무관질문모형들의 중복시행

이 장에서는 Greenberg et al.의 무관질문모형과 2단계 무관질문모형, 그리고 개선된 무관질문모형을 중복시행모형으로 확장하여 각 모형들을 m번 중복시행한 후 그 때까지 “예”라고 응답한 회수를 가지고 민감한 속성에 대한 모비율 π를 추정하고자 한다.

3.1 Greenberg et al.의 무관질문모형의 중복시행

Greenberg et al.의 무관질문모형을 m번 중복시행했을 때, j번째 응답자가 i번 “예”라고 응답한 것을 $X_j = i$로 나타내면 j번째 응답자가 i번 “예”라고 응답할 확률은 다음과 같다.

$$\begin{aligned} P(X_j = i \mid m) &= \binom{m}{i}\left[\pi\{p+(1-p)\pi_y\}^i\{(1-p)(1-\pi_y)\}^{m-i} + (1-\pi)\{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}\right] \\ &= w_i, \quad i=0,1,2,\cdots,m,\ j=1,2,\cdots,n. \end{aligned} \tag{3.1}$$

여기서, $\sum_{i=0}^{m} w_i = 1$이다.

단순임의복원추출된 n명의 응답자들 중에서 i번 “예”라고 응답한 사람의 수를 n_i라고 하면 $\sum_{i=0}^{m} n_i = n$이고, 이 때 우도함수와 로그 우도함수는 다음과 같다.

$$L = \prod_{i=0}^{m} w_i^{\,n_i}, \quad \log L = \sum_{i=0}^{m} n_i \log w_i . \tag{3.2}$$

민감한 속성에 대한 모비율 π의 최대우도추정량 $\hat{\pi}_{u_1}$를 구하기 위하여 방정식 $\frac{\partial \log L}{\partial \pi} = 0$를 직접 계산하는 것은 이 방정식이 폐쇄형으로 표현되지 않기 때문에 쉽지 않다. 그러나, 이러한 경우 반복적인 스코링 방법(scoring method)을 이용하여 $\hat{\pi}_{u_1}$를 구할 수 있는 데 그 방법을 간략히 살펴보면 다음과 같다.

우선, π의 임의의 초기치 π_0 $(0 \le \pi_0 \le 1)$를 선택하여 π의 첫 번째 근사치 π_1을 다음과 같이 구한다.

$$\pi_1 = \pi_0 + \frac{S(\pi_0)}{I(\pi_0)} = \pi_0 + \delta_0 \,. \tag{3.3}$$

여기서, $S(\pi_0)$와 $I(\pi_0)$는 각각 π_0에서의 스코어(score)와 정보(information)이며, δ_0는 수정항이다.

또 다시 식(3.3)의 우변 항에 있는 π_0 대신에 π_1을 대입하여 π의 두 번째 근사치 π_2를 구한다. 이러한 과정을 수정항이 무시될 수 있을 정도로 작아질 때까지 반복적으로 계속하여 π의 추정량 $\hat{\pi}_{u_1}$를 구한다. 이렇게 얻어진 추정량 $\hat{\pi}_{u_1}$는 편향이 매우 작아지게 되어 점근적으로 비편향추정량이 되며, $\hat{\pi}_{u_1}$의 분산은 다음과 같이 점근적으로 정보의 역수가 된다.

$$V_1(\hat{\pi}_{u_1}) \approx \frac{1}{I_1(\pi)} \,. \tag{3.4}$$

이 때, Greenberg et al.의 무관질문모형의 중복시행에서 스코어 $S_1(\pi)$와 정보 $I_1(\pi)$를 유도해 보면 다음과 같다.

$$\begin{aligned} S_1(\pi) &= \frac{\partial \log L}{\partial \pi} = \sum_{i=0}^{m} \frac{n_i}{w_i}\left(\frac{\partial w_i}{\partial \pi}\right) \\ &= \sum_{i=0}^{m} \frac{n_i[\{p+(1-p)\pi_y\}^i\{(1-p)(1-\pi_y)\}^{m-i} - \{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}]}{\pi\{p+(1-p)\pi_y\}^i\{(1-p)(1-\pi_y)\}^{m-i} + (1-\pi)\{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}} \,, \end{aligned} \tag{3.5}$$

$$\begin{aligned} I_1(\pi) &= -E\left(\frac{\partial^2 \log L}{\partial \pi^2}\right) = n\sum_{i=1}^{m} \frac{1}{w_i}\left(\frac{\partial w_i}{\partial \pi}\right)^2 \\ &= n\sum_{i=0}^{m} \frac{\binom{m}{i}[\{p+(1-p)\pi_y\}^i\{(1-p)(1-\pi_y)\}^{m-i} - \{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}]^2}{\pi\{p+(1-p)\pi_y\}^i\{(1-p)(1-\pi_y)\}^{m-i} + (1-\pi)\{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}} \,. \end{aligned} \tag{3.6}$$

3.2 2단계 무관질문모형의 중복시행

2단계 무관질문모형을 m번 중복시행했을 때, j번째 응답자가 i번 “예”라고 응답

한 것을 $X_j = i$로 나타내면 j번째 응답자가 i번 “예”라고 응답할 확률은 다음과 같다.

$$P(X_j = i \mid m) = \binom{m}{i}\Big[\pi\{T + (1-T)(p+(1-p)\pi_y)\}^i\{(1-T)(1-p)(1-\pi_y)\}^{m-i} + (1-\pi)\{(1-T)(1-p)\pi_y\}^i\{1-(1-T)(1-p)\pi_y\}^{m-i}\Big] = w_i, \quad i = 0,1,2,\cdots,m, \; j = 1,2,\cdots,n. \tag{3.7}$$

여기서, $\sum_{i=0}^{m} w_i = 1$이다.

2단계 무관질문모형의 중복시행에서 반복적인 스코링 방법을 이용하여 π의 최대우도추정량 $\hat{\pi}_{u_2}$를 구할 수 있다. 이렇게 얻은 $\hat{\pi}_{u_2}$는 점근적으로 비편향추정량이 되며, $\hat{\pi}_{u_2}$의 분산 $V_2(\hat{\pi}_{u_2})$는 점근적으로 정보 $I_2(\pi)$의 역수가 된다. 이 때, 스코어 $S_2(\pi)$와 정보 $I_2(\pi)$를 유도해 보면 다음과 같다.

$$S_2(\pi) = \sum_{i=0}^{m} \frac{n_i[A^i(1-A)^{m-i} - B^i(1-B)^{m-i}]}{\pi A^i(1-A)^{m-i} + (1-\pi)B^i(1-B)^{m-i}}, \tag{3.8}$$

$$I_2(\pi) = n\sum_{i=0}^{m} \frac{\binom{m}{i}[A^i(1-A)^{m-i} - B^i(1-B)^{m-i}]^2}{\pi A^i(1-A)^{m-i} + (1-\pi)B^i(1-B)^{m-i}}. \tag{3.9}$$

여기서 $A = T + (1-T)\{p+(1-p)\pi_y\}$, $B = (1-T)(1-p)\pi_y$이다.

3.3 개선된 무관질문모형의 중복시행

개선된 무관질문모형을 m번 중복시행했을 때, j번째 응답자가 i번 “예”라고 응답한 것을 $X_j = i$로 나타내면 j번째 응답자가 i번 “예”라고 응답할 확률은 다음과 같다.

$$P(X_j = i \mid m) = \binom{m}{i}(1-\pi)\{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i} = w_i, \quad i < m, \; j = 1,2,\cdots,n. \tag{3.10}$$

$$P(X_j = m \mid m) = \pi + (1-\pi)\{(1-p)\pi_y\}^m = w_m, \quad j = 1,2,\cdots,n. \tag{3.11}$$

여기서, $\sum_{i=0}^{m} w_i = 1$이다.

개선된 무관질문모형의 중복시행에서 반복적인 스코링 방법을 이용하여 π의 최대우도추정량 $\hat{\pi}_{u_3}$를 구할 수 있다. 이렇게 얻은 $\hat{\pi}_{u_3}$는 점근적으로 비편향추정량이 되며, $\hat{\pi}_{u_3}$의 분산 $V_3(\hat{\pi}_{u_3})$는 점근적으로 정보 $I_3(\pi)$의 역수가 된다. 이 때, 스

코어 $S_3(\pi)$와 정보 $I_3(\pi)$를 유도해 보면 다음과 같다.

$$S_3(\pi) = -\sum_{i=0}^{m-1} \frac{n_i\{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}}{1-\pi} + n_m \frac{1-\{(1-p)\pi_y\}^m}{\pi+(1-\pi)\{(1-p)\pi_y\}^m}, \quad (3.12)$$

$$I_3(\pi) = n\left[\sum_{i=0}^{m-1} \frac{\binom{m}{i}\{(1-p)\pi_y\}^i\{1-(1-p)\pi_y\}^{m-i}}{1-\pi} + \frac{[1-\{(1-p)\pi_y\}^m]^2}{\pi+(1-\pi)\{(1-p)\pi_y\}^m}\right]. \quad (3.13)$$

한편, 제안한 중복시행모형들에 대하여 스코링 방법을 이용하여 모비율의 추정치와 그 수정항에 대한 모의실험을 해 본 결과, 그리 많지 않은 반복회수에 수정항이 무시될 수 있을 정도로 작아짐을 알 수 있었다.

4. 중복시행 무관질문모형들의 효율성 비교

이 장에서는 Liu-Chow의 중복시행모형과 3장에서 다룬 중복시행 무관질문모형들과의 효율성을 비교해 보고자 한다.

Liu-Chow의 모형은 Warner의 모형을 m번 중복시행한 모형이며, j번째 응답자가 i번 "예"라고 응답한 것을 $X_j = i$로 나타내면 j번째 응답자가 i번 "예"라고 응답할 확률은 다음과 같다.

$$P(X_j = i \mid m) = \binom{m}{i}[\pi p^i(1-p)^{m-i} + (1-\pi)p^{m-i}(1-p)^i] = w_i, \quad (4.1)$$
$$i = 0, 1, 2, \cdots, m, \; j = 1, 2, \cdots, n.$$

Liu-Chow의 중복시행모형에서 얻어진 추정량 $\hat{\pi}_{lc}$는 편향이 매우 작아지게 되어 점근적으로 비편향추정량이 되며, $\hat{\pi}_{lc}$의 분산은 다음과 같이 점근적으로 정보의 역수가 된다.

$$V(\hat{\pi}_{lc}) \approx \frac{1}{I(\pi)}. \quad (4.2)$$

이 때, 정보 $I(\pi)$는 다음과 같다.

$$I(\pi) = n\sum_{i=0}^{m} \frac{\binom{m}{i}[p^i(1-p)^{m-i} - p^{m-i}(1-p)^i]^2}{\pi p^i(1-p)^{m-i} + (1-\pi)p^{m-i}(1-p)^i}. \quad (4.3)$$

Liu-Chow의 중복시행모형에서의 분산과 3장에서 다룬 중복시행 무관질문모형들의 분산을 이용하여 중복시행의 수 m이 2일 때, 효율성을 수치적으로 비교하고자 한다. $n = 100$, $T = 0.2$일 때 π와 π_y 및 p를 변화시켜 가면서 분산비 $V_1 = V(\hat{\pi}_{lc})$

/ $V_1(\hat{\pi}_{u_1})$, $V_2 = V(\hat{\pi}_{lc})/V_2(\hat{\pi}_{u_2})$, $V_3 = V(\hat{\pi}_{lc})/V_3(\hat{\pi}_{u_3})$를 계산한 결과 다음 <표 4.1>을 얻었다. 이 때, π값은 민감한 속성의 모비율이므로 비교적 작은 값을 가질 것이라고 생각하여 0.1에서 0.5까지 0.1씩 증가시켰으며, 무관한 속성의 모비율 π_y와 민감한 속성이 선택될 확률 p는 0.2에서 0.8까지 0.2씩 증가시켜가면서 효율성의 변화를 살펴보았다.

<표 4.1> 제안한 중복시행 무관질문모형들과 Liu-Chow의 중복시행모형과의 효율성 비교

π	π_y	0.2			0.4			0.6			0.8		
	p	V_1	V_2	V_3	V_1	V_2	V_3	V_1	V_2	V_3	V_1	V_2	V_3
0.1	0.2	0.13	0.45	2.08	0.08	0.32	1.22	0.07	0.27	0.65	0.08	0.28	0.33
	0.4	7.04	11.77	28.52	5.10	9.44	20.29	4.41	8.31	13.13	4.36	7.92	8.18
	0.6	15.26	18.85	30.71	13.12	17.22	25.89	11.88	16.19	20.29	11.28	15.53	15.27
	0.8	1.95	2.09	2.59	1.90	2.06	2.47	1.88	2.07	2.29	1.87	2.08	2.08
0.2	0.2	0.16	0.49	1.87	0.11	0.38	1.35	0.10	0.35	0.85	0.12	0.37	0.47
	0.4	5.39	8.26	17.92	4.31	7.14	14.72	4.00	6.80	11.02	4.18	6.95	7.70
	0.6	10.24	12.20	18.63	9.23	11.38	16.99	8.99	11.32	14.72	9.17	11.61	12.24
	0.8	1.65	1.75	2.10	1.60	1.71	2.05	1.63	1.74	1.97	1.68	1.80	1.87
0.3	0.2	0.18	0.51	1.81	0.13	0.42	1.42	0.13	0.40	0.98	0.15	0.44	0.59
	0.4	4.70	7.00	14.40	3.92	6.17	12.55	3.77	6.05	10.09	4.09	6.44	7.56
	0.6	8.55	10.05	14.79	7.76	9.35	13.89	7.72	9.40	12.55	8.17	9.93	10.94
	0.8	1.58	1.66	1.95	1.51	1.61	1.92	1.54	1.63	1.87	1.61	1.70	1.81
0.4	0.2	0.19	0.52	1.78	0.14	0.43	1.48	0.14	0.42	1.09	0.17	0.48	0.69
	0.4	4.35	6.44	12.98	3.71	5.73	11.67	3.65	5.70	9.80	4.09	6.22	7.69
	0.6	7.83	9.18	13.24	7.13	8.51	12.63	7.13	8.55	11.67	7.70	9.16	10.47
	0.8	1.56	1.64	1.89	1.48	1.57	1.87	1.49	1.58	1.83	1.57	1.66	1.78
0.5	0.2	0.18	0.51	1.79	0.14	0.44	1.53	0.14	0.44	1.18	0.18	0.51	0.79
	0.4	4.16	6.21	12.63	3.64	5.60	11.58	3.64	5.60	10.01	4.16	6.21	8.13
	0.6	7.60	8.93	12.83	6.95	8.29	12.35	6.95	8.29	11.58	7.60	8.93	10.58
	0.8	1.56	1.64	1.88	1.48	1.57	1.86	1.48	1.57	1.83	1.56	1.64	1.79

<표 4.1>에서 1보다 큰 값은 3장에서 다룬 중복시행 무관질문모형들이 Liu-Chow의 중복시행모형보다 더 효율적임을 나타낸다. 중복시행 무관질문모형들은 p값이 0.4보다 클 때, π와 π_y값의 크기에 관계없이 Liu-Chow의 중복시행모형보다 효율이 좋게 나타나고 있다. 특히, p값이 0.5에 가까울 때, 각 중복시행 무관질문모형들의 효율이 더 좋아짐을 알 수 있다.

중복시행 무관질문모형들 간의 비교에서는 개선된 무관질문모형의 효율이 가장 좋고, 그 다음으로 2단계 무관질문모형 그리고, Greenberg et al.의 무관질문모형 순으로

효율이 좋게 나타나고 있다. 그리고, 중복시행 개선된 무관질문모형은 π_y 값이 클수록 효율적임을 알 수 있다.

한편, T 값을 변화시켜 가면서 효율성을 비교해 본 결과 전반적으로 <표 4.1>과 비슷한 경향을 나타내고 있었으나, T 값이 클수록 중복시행 2단계 무관질문모형의 효율이 좋게 나타나는 결과를 얻을 수 있었다. 이러한 결과는 T 값이 1에 가까워질수록 직접질문의 형태를 취하기 때문에 그 만큼 분산이 작아지기 때문인 것으로 분석된다. 극단적으로 T 값이 1인 경우에는 직접질문 형태와 일치하게 되어 분산이 가장 작게 나타나게 된다.

또한, 중복시행의 수 m 을 크게 증가시킬수록 각 무관질문모형들의 효율이 좋아지는 결과를 얻을 수 있었다.

5. 결론

Liu-Chow는 확률화응답모형에서 추정량의 분산이 커지는 단점을 해결하기 위하여 Warner의 모형을 중복하여 사용함으로서 표본의 크기를 증가시키지 않고서도 추정량의 분산을 줄일 수 있는 중복시행모형을 제안하였다.

본 논문에서는 Greenberg et al.의 무관질문모형과 김종호외 2인의 2단계 무관질문모형, 그리고 이기성과 홍기학의 개선된 무관질문모형을 중복시행모형으로 확장하여 Liu-Chow의 중복시행모형과 그 효율성을 비교하였다. 그 결과 중복시행 무관질문모형들은 p 값이 0.4보다 클 때, π와 π_y 값의 크기에 관계없이 Liu-Chow의 중복시행모형보다 효율이 좋게 나타나고 있었으며, 특히, p 값이 0.5에 가까울 때, 각 중복시행 무관질문모형들의 효율이 더 좋아짐을 알 수 있었다. 또한, 중복시행 무관질문모형들 간의 비교에서는 개선된 무관질문모형의 효율이 가장 좋았고, 그 다음으로 2단계 무관질문모형 그리고, Greenberg et al.의 무관질문모형 순으로 효율이 좋게 나타나고 있었으며 중복시행 개선된 무관질문모형은 π_y 값이 클수록 더욱 더 효율적이었다.

참고문헌

1. 김종호, 류제복, 이기성 (1992). 새로운 2단계 확률화응답모형, 응용통계연구, 제5권 2호, 157-167.
2. 류제복, 홍기학, 이기성 (1993). 확률화응답모형, 자유아카데미, 서울.
3. 이기성, 홍기학 (1998). 개선된 무관질문모형, 응용통계연구, 제11권 2호, 415-421.
4. 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.
5. Liu, P. T. and Chow, L. P. (1976). The Efficiency of the Multiple Trial Randomized Response Technique, *Biometrics*, 32, 607-618.
6. Mangat, N. S. (1994). An Improved Randomized Response Strategy, *Journal*

of the Royal Statistical Society : Series B, 56, 93-95.
7. Mangat, N. S. and Singh, R. (1990). An Alternative Randomized Response Procedure, *Biometrika*, 77, 439-442.
8. Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, John Wiley and Sons, New York.
9. Warner, S. L. (1965). Randomized Response ; A Survey Technique for Eliminating Evasive Answer Bias, *Journal of the American Statistical Association*, 60, 63-69.