Proceedings of Joint Conference of Korean Data And Information Science Society
and The Korean Data Analysis Society,
April 28 - 29, 2006, pp. 25-32

청렴도 조사에서 간접질문에 의한 청렴도 점수의 계산

남기성[1]

요 약

본 논문에서는 민감한 사항 특히 청렴도 조사에서 금품·향응제공여부 등과 같이 응답자들이 정직하게 응답하지 않을 수 있는 질문들에 대하여 응답자의 비밀을 노출시키지 않고서 양질의 보다 정확한 정보를 얻을 수 있는 확률화응답기법을 적용하여 실제 청렴도 점수를 추정하고자 한다. 본 연구는 청렴도 조사와 같은 민감한 질문에 대하여 응답자에게 심리적 안정감을 주어 거짓응답을 줄이면서 보다 진실에 가까운 청렴도 점수를 구할 수 있는 사례를 제시한다.

주요용어 : 청렴도조사, 확률화응답기법, 민감한 질문, 청렴도점수, 추정.

1. 서론

청렴(淸廉, integrity)이란 마음이 고결하고 재물 욕심이 없음을 뜻하며, 반대로 부패(腐敗, rotting)는 썩음, 정신적으로 타락함을 의미하는 말이다.

부패문제를 근절하기 위해서는 부패의 본질을 규명하고 이를 해소하기 위해 광범위하고 지속적인 특단의 조치가 필요하며, 부패 본질에 대한 규명을 위해서는 무엇보다 공직부패의 발생현황과 실태분석을 통해 부패유형, 부패수준 및 추이를 파악하는 등 객관적이고 포괄적인 분석이 선행되어야 한다. 또한, 부패발생을 근본적으로 방지하기 위해서는 비리공직자에 대한 적발과 처벌에 의존하는 사후 통제방법을 벗어나 새로운 접근방법으로서 부패발생 요인을 사전에 제거하는 예방적인 반부패 정책이 요구된다. 예방적 반부패 전략을 마련하기 위해서는 행정기관 또는 업무 분야별로 부패를 유발하는 요인들의 상황, 즉 요인이 성숙되어 부패가 유발될 가능성에 대한 규명과 분석이 전제되어야 한다. 이를 위해서는 부패방지 노력이 객관적인 활동지표와 더불어 결과측면에서도 평가되어야 하는데 이는 부패유발 요인들이 어떻게 작용하며 변화되고 있는가를 계량화하는 작업이 될 것이다. 따라서 우리나라의 부패방지위원회는 공공기관의 자율적인 반부패 노력을 유도하고 부패유발요인을 도출하여 이에 대한 제도를 개선함으로써 부패를 방지하는 과학적이고 체계적인 접근을 위해 공공기관에 대한 청렴도 측정을 추진하게 되었다(2003, 부패방지백서).

청렴도 조사에는 금품 향응 제공에 대한 매우 민감한 질문이 포함되어있다. 청렴도 조사 과정에서 많은 설득과 법적인 비밀보장에 대한 홍보를 하면서 조사를 수행하나 향후의 불이익 등에 대한

[1]부산광역시 금정구 장전동 산30번지 부산대학교, 사회조사연구소 연구원.
E-mail : namkiseong@hanafos.com

염려로 거짓응답으로 일관하는 경우가 흔히 있다. 만일 청렴도 조사에서 매우 민감한 문제에 대하여 거짓으로 응답한다면 체감 청렴도는 매우 높아지며, 조사의 존재 자체가 흔들리게 된다.

따라서 본 연구는 통계학 분야에서 사용하는 특수한 조사 기법인 확률화응답기법(randomized response technique ; RRT)을 청렴도 조사와 같이 거짓응답을 할 가능성이 매우 높은 조사에 적용하고자 하는 것이다.

RRT는 거짓응답과 같은 문제점을 해결하기 위하여 Warner(1965)가 제안한 확률장치를 이용한 간접질문방식으로 인터넷 혹은 전자우편, 전화조사를 이용한 설문조사에 적용해 볼 수 있다. 이때, 응답자들은 인터넷상 혹은 미리 준비된 확률장치를 통해 선택된 설문에 대하여 응답을 하게 되므로 응답자 자신의 비밀이나 개인 정보의 유출을 이유로 정확하지 않은 응답을 할 가능성을 줄일 수 있게 된다. 물론, 응답자들에게 RRT를 이해시켜야 되므로 응답자들의 관심을 불러일으키지 못할 수 있다는 문제점이 있다. 하지만, 확률화응답기법을 보다 쉽게 이해할 수 있도록 충분히 설명하고, 다양한 형태의 확률장치를 구현하여 응답자들의 흥미를 유발시킨다면 이러한 문제점은 해결될 수 있으리라 생각된다.

본 논문은 1장 서론에서 청렴도조사에서 거짓응답을 줄이기 위한 조사기법으로 확률화응답기법의 필요성을 설명하고, 2장에서는 청렴도 조사와 조사의 문제점에 대하여 살펴보고, 3장에서는 확률화응답기법에서의 모수의 추정에 대하여 알아보고, 4장에서는 확률화응답기법을 이용한 청렴도 점수의 추정 및 모의실험을 하고, 5장에서는 결론을 다룬다.

2. 청렴도 조사

청렴도 측정을 위한 조사는 반부패특별위원회에서 1999년도에 청렴도 측정모형을 마련하여 측정모형의 적합성 검증을 위한 시험측정을 2000년과 2001년 총 3회에 걸쳐 공공기관을 대상으로 실시하였다. 1회 측정은 측정모형의 적정성을 확인하기 위해 실시하였으며, 그 결과를 토대로 측정모형을 보완하고 측정대상기관을 추가 확대하여 2차·3차 시험측정을 실시하였다.

2002년 1월 25일 부패방지위원회 출범 후 시험측정 결과를 토대로 모형의 통계적 타당성에 대한 전문가 의견을 최종적으로 수렴하여 측정모형을 완성하고 71개 공공기관의 348개 업무에 대해 청렴도 측정을 실시하여 2003년 3월 31일 부패방지대책 대국민보고회 개최 후 4월 10일 기관별 청렴도 점수를 공개하였다.

2003년 청렴도 측정은 2002년 측정결과 공개 후 제기된 의견을 반영하여 조사표를 수정하는 등 측정모형을 일부 수정하였으나 청렴도 측정결과 후 기관별로 청렴도가 개선되는 정도를 분석하기 위해 동일한 모형을 통해 77개 기관의 394개 업무를 대상으로 기관별 청렴도를 측정하였다.

청렴도 조사는 <표 2.1>의 체감청렴도와 잠재청렴도의 합인 종합청렴도로 나타내며, 일반적으로 10점 혹은 100점 만점으로 계산되어진다. 여기서 많이 행해지는 거짓응답은 금품·향응제공 경험이 있는데도 불구하고 경험이 없는 것으로 거짓응답을 하는 것이다. 그러면 부패경험의 하위측도

는 만점으로 나오며, 전체적인 종합청렴도는 매우 크게 증가하게 된다.

<표 2.1> 청렴도 평가영역 및 영역별 가중치

상위영역	하위영역	설문항목
체감청렴도 (0.494)	부패경험(0.483)	금품·향응 제공빈도(0.544)
		금품·향응 제공규모(0.456)
	부패인식(0.517)	금품·향응 제공인식(1.000)
잠재청렴도 (0.506)	업무환경(0.241)	금품·향응 제공 관행화(0.667)
		추가면담 필요성(0.333)
	행정제도(0.237)	기준·절차의 현실성(0.569)
		정보공개 정도(0.431)
	개인태도(0.294)	업무처리 공정성(0.599)
		금품·향응 수수기대(0.401)
	부패통제(0.228)	부패방지 노력도(0.585)
		이의제기 용이성(0.415)

<표 2.2>는 직접 설문조사한 국영기업체 2곳의 조사결과이다. <표 2.2>에서와 같이 '업무처리 시 직원에게 금품이나 향응을 제공하는 것이 관행으로 굳어져 거부하기가 어려웠다고 생각하십니까?'에 대하여 1.4%와 0.3%가 긍정적인 응답을 하였으며, '업무처리 과정에서 업무처리 대가로 금품이나 향응을 제공할 경우 이를 받았을 것이라고 생각하십니까?'에 대하여 0.4%와 3.5%가 긍정적인 응답을 하였고, '업무처리 경험에 비추어 볼 때, 업무 처리 시 금품이나 향응이 제공되고 있었다고 생각하십니까?'에 대하여 1.6%와 7.0%가 긍정적인 응답을 하여 두 회사 모두 금품 향응의 제공 징후가 포착되고 있다. 그러나 직접적인 금품향응 경험에 대하여서는 0%와 0.2%로 제공을 하지 않

<표 2.2> 청렴도 설문조사 결과

문항	항목	A기업체		B기업체	
		빈도	백분율	빈도	백분율
금품 향응 제공 관행화 업무처리시 직원에게 금품이나 향응을 제공하는 것이 관행으로 굳어져 거부하기가 어려웠다고 생각하십니까?	매우 그렇다	4	.8	3	.3
	그런 편이다	3	.6	-	-
	보통이다	24	4.6	24	2.2
	그렇지 않은 편이다	152	29.2	255	23.4
	전혀 그렇지 않다	338	64.9	807	74.1
금품향응기대 업무처리 과정에서 업무처리 대가로 금품이나 향응을 제공할 경우 이를 받았을 것이라고 생각하십니까?	매우 그렇다	-	-	8	.7
	그런 편이다	2	.4	31	2.8
	보통이다	16	3.1	30	2.8
	그렇지 않은 편이다	115	22.1	332	30.5
	전혀 그렇지 않다	388	74.5	688	63.2
금품향응제공인식 업무처리 경험에 비추어 볼 때, 업무처리시 금품이나 향응이 제공되고 있었다고 생각하십니까?	매우 그렇다	2	.4	34	3.1
	그런 편이다	6	1.2	43	3.9
	보통이다	22	4.2	35	3.2
	그렇지 않은 편이다	164	31.5	238	21.9
	전혀 그렇지 않다	327	62.8	739	67.9
금품이나 향응을 제공한 경험	있다	-	-	2	.2
	없다	521	100.0	1087	99.8
합계		521	100.0	1089	100.0

고 있다는 식으로 긍정적인 응답 비율이 아주 낮게 나타난다. 이는 이 답변에 대하여 불이익이나 또 다른 피해를 볼 수 있다는 생각에서 거짓 응답을 하고 있는 것이다.

3. 확률화응답기법

사회 여러 분야의 조사에는 응답자들이 응답을 회피하거나 정직하게 응답하지 않는 질문들이 종종 포함된다. 예를 들면, 불로소득, 탈세여부, 전과 경험, 컨닝 경험, 음주운전경험, 알코올중독, 환각제사용, 성경험, 낙태경험, 동성연애 등과 같은 사항들은 응답자들이 응답을 꺼려하는 질문들이다. 이러한 민감한 질문을 공개적으로 하게 되면 무응답이나 거짓응답 또는 응답을 회피함으로써 응답자들로부터 정확한 정보를 얻을 수 없게 된다.

따라서 민감한 질문에 대해 보다 신뢰할 수 있는 정보를 얻기 위해서는 직접질문보다는 간접적인 대체 질문방식이 필요하게 된다. 이에 Warner(1965)는 응답자들에게 민감한 질문과, 민감한 질문과 배반되는 즉, 부의 관계를 갖는 질문 중에서 확률장치를 통해 선택된 질문에 응답하게 함으로써 응답자의 신분이나 비밀을 노출시키지 않고서 민감한 질문에 대한 정보를 이끌어 낼 수 있는 확률화응답기법(randomized response technique ; RRT)을 처음으로 제시하였다.

응답자들은 확률장치에 의해 선택된 질문에 응답하면 되며, 이 때 조사자는 응답자가 어떤 질문에 응답을 했는지를 알 수 없게 되므로 응답자는 솔직하게 응답할 수 있다. Warner가 사용한 관련 질문기법의 확률장치는 다음과 같은 2개의 설문으로 적용할 수 있다.

설문 1 : 귀하는 00업체와 계약관계로 금품이나 향응을 제공한 경험이 있습니까?
설문 2 : 귀하는 00업체와 계약관계로 금품이나 향응을 제공한 경험이 없습니까?

이 때, 단순임의복원으로 추출된 n명의 응답자들은 확률장치에 의해서 선택된 설문에 따라 "예" 또는 "아니오"로만 응답한다. 설문 1이 선택될 확률을 p, 설문 2가 선택될 확률을 $1-p$라 하면 응답자들이 "예"라고 응답할 확률은 다음과 같다.

$$\lambda = p\pi + (1-p)(1-\pi) \tag{2.1}$$

여기서, π는 민감한 속성 A에 대한 모비율이다.

n명의 응답자들 중에서 "예"라고 응답한 사람들의 수를 n'라고 하면, $\hat{\lambda} = n'/n$이 된다. 따라서 민감한 속성 A에 대한 모비율 π의 추정량은 다음과 같다.

$$\hat{\pi}_w = \frac{\hat{\lambda}-(1-p)}{2p-1}\ ,\ p \neq \frac{1}{2} \tag{2.2}$$

다음으로 양적 정보를 얻기 위한 Greenberg 등(1971)의 무관질문기법이 있는데 이 방법을 이용하기 위한 확률장치로 다음과 같은 두 가지 설문을 생각할 수 있다.

설문 1 : 귀하가 00업체와 계약관계로 금품이나 향응을 제공한 금액은 얼마입니까?
설문 2 : 우리나라 국민이 한 달에 사용하는 용돈의 평균은 얼마정도라고 생각하십니까?

이 때, 설문 1의 금품이나 향응을 제공한 금액과 같은 민감한 변수 X에 대한 모평균을 μ_x라 하고, 설문 2와 같이 민감한 변수와 무관한 변수 Y에 대한 모평균을 μ_y라 하자. μ_y를 모르는 경우에 구하고자하는 미지의 모수가 μ_x와 μ_y이므로 최소한 두개의 표본이 필요하며, 모집단으로부터 단순임의복원으로 크기가 n_1과 n_2인 두 개의 독립 표본을 추출한다. 두 개의 표본을 사용해야 하므로 $i(i=1,2)$번째 표본에서 민감한 설문이 선택될 확률이 p_i가 되는 두 개의 확률장치가 필요하게 된다.

$i\ (i=1,2)$번째 표본의 응답자들이 Z_i라고 응답할 경우 그 평균은 다음과 같다.

$$\mu_{z_i} = p_i\mu_x + (1-p_i)\mu_y \tag{2.3}$$

n_i명의 응답자들이 확률장치에 의해서 선택된 설문에 대하여 $z_{ij}(i=1,2, j=1,2,\cdots,n_i)$라고 응답했을 때, 민감한 변수 X에 대한 모평균 μ_x의 추정량 $\hat{\mu}_x$는 다음과 같다.

$$\hat{\mu}_x = \frac{(1-p_2)\bar{z}_1 - (1-p_1)\bar{z}_2}{p_1-p_2}\ ,\ \text{단}\quad p_1 \neq p_2 \tag{2.4}$$

여기서, $\bar{z}_1 = (1/n_1)\sum_{j=1}^{n_1} z_{1j}$, $\bar{z}_2 = (1/n_2)\sum_{j=1}^{n_2} z_{2j}$이다.

미국, 캐나다, 영국, 호주 등 서구 여러 나라와 일본, 인도 등 몇몇 아시아 국가에서도 이 분야에 대한 연구가 활발히 진행되고 있다. 특히, Fox와 Tracy(1986), Chaudhuri와 Mukerjee(1988)는 확률화응답기법들을 정리, 요약하여 체계화하였다. 최근에는 이들 이론들의 실제적 활용에 많은 관심이 집중되고 있으며, 사회학, 경영학, 의학 등 여러 학문분야에서의 조사활동에도 이들 방법의 활용이 적극 모색되고 있다. 또한 남기성, 이기성 등(2001, 2002)은 인터넷에서 RRT를 사용할 수 있도록 시스템을 개발하여 발표하였다.

4. 확률화응답기법을 적용한 청렴도 점수의 추정

청렴도 조사에서 RRT 기법을 적용하는데 조사방법으로 사용할 수 있는 것은 전화조사, 인터넷조사, 전자우편조사의 3 가지로 나누어 볼 수 있으며, 전화조사는 비용과 시간 등이 적게 든다는 장점이 있으나, 확률장치를 설명하기 어려운 단점이 존재하며, 인터넷조사는 응답자가 조사기간에 직접 인터넷에 접속하여야하는 문제로 능동적인 조사가 어려운 단점이 존재한다. 따라서 청렴도 조사에서 RRT를 적용하기 위해서는 전자우편조사가 가장 적합하다. 즉, 사전에 준비된 메일주소로

조사의 목적과 RRT를 소개한 글과 RRT가 HTML 기법으로 구성된 전자우편을 발송하고, 응답자는 응답 후 보내기만 시행하면 응답결과가 조사자에게 발송되도록 시스템을 구성하는 것이다. 사전에 전화상으로 조사 시기와 목적, 방법을 알려준다면 보다 응답 성공률을 높일 수 있다.

청렴도 조사에서 RRT가 필요한 문항은 금품·향응 제공 여부와 금품·향응 제공 빈도 및 규모에서 RRT를 적용하여야 거짓응답을 줄일 수 있다. 청렴도 조사 문항과 관련한 RRT는 <표 4.1>과 같다.

<표 4.1> 청렴도 조사 문항과 관련한 RRT

청렴도 조사 문항	RRT 기법
직원에게 업무처리와 관련하여 금품이나 향응을 제공한 경험이 있습니까?	질적자료 관련질문기법
몇 번 정도 금품이나 향응을 제공했습니까?	양적 무관질문기법
1년 동안 금품이나 향응을 제공한 규모는 얼마정도 됩니까?	양적 무관질문기법
제공한 금품이나 향응은 어떤 종류였습니까?	다지질문기법
금품이나 향응을 제공한 이유는 무엇입니까?	다지질문기법

다음으로 부패방지위원회에서 사용하는 청렴도조사에서의 청렴도의 점수는 식 (4.1)과 같다.

$$\text{부패경험} = 0.483(0.544\ \hat{\mu}_{x1} + 0.456\ \hat{\mu}_{x2}) \tag{4.1}$$

여기서, 0.483은 <표 2.1>에서의 부패경험의 가중치이고, 0.544는 금품·향응 제공빈도의 가중치이며, 0.456은 금품·향응 제공규모의 가중치이다. 그리고 $\hat{\mu}_{x1}$과 $\hat{\mu}_{x2}$는 다음과 같이 산출한다.

$$\text{기관별 금품제공빈도 점수}(\ \hat{\mu}_{x1}) = 10\times(1-\frac{\text{기관별 1인 제공빈도}}{UCP})$$

$$\text{기관별 금품제공규모 점수}(\ \hat{\mu}_{x2}) = 10\times(1-\frac{\text{기관별 1인 제공규모}}{UCP})$$

단, UCP(Upper Cut of Point)는 전체기관 평균(μ) 및 표준편차(σ)를 이용하여 $\mu+2.0\sigma$에 해당하는 값이다.

따라서 본 논문에서는 직접질문에 해당하는 기관별 1인 제공빈도 대신 RRT를 이용하여 구한 $\hat{\pi}_w$를 이용하여 기관별 금품제공빈도 점수와 기관별 금품제공규모 점수를 대체하는 것을 제안하고자 한다. <표 4.2>는 RRT기법을 이용하여 부패 경험비율($\hat{\pi}_w$)을 산출하는 과정이다. 예를 들어 어떤 기관에 1,000명의 임직원이 있을 때 각각의 예상 금품 제공 인원수에 따른 계산 결과이다. 여기서 전체 인원은 1,000명으로 설문1이 나올 확률 p=0.3으로 고정하였다.(단, 결과는 확률 이론적인 예상 비율임). 양적 무관질문기법을 이용하는 기관별 금품제공규모 점수($\hat{\mu}_{x2}$)는 생략하였다.

<표 4.2> 모의실험 결과

전체 인원(n)	1,000	1,000	1,000	1,000	1,000
금품 제공자 예상 수	30	50	100	150	200
설문1이 선택될 확률(p)	0.3	0.3	0.3	0.3	0.3
설문1에 배당될 인원수	300	300	300	300	300
설문2에 배당될 인원수	700	700	700	700	700
설문1에서 "예" 응답자	9	15	30	45	60
설문2에서 "예" 응답자	679	665	630	595	560
전체에서 "예" 응답자(n')	688	680	660	640	620
$\hat{\lambda} = \frac{n'}{n}$	0.69	0.68	0.66	0.64	0.62
$\hat{\pi}_w$	0.03	0.05	0.1	0.15	0.2

5. 결론

본 논문에서는 민감한 정보 특히 청렴도 조사에서와 같이 응답자들이 정직하게 응답하기를 꺼리는 질문들에 대하여 직접응답 대신에 간접응답을 통해 응답자의 비밀을 노출시키지 않고서 보다 정확한 정보를 얻을 수 있는 간접응답기법인 확률화응답기법을 적용하는 것을 제안하였으며 청렴도 점수를 추정하기 위한 식을 제시하였다.

이는 기업체 등에서 개인의 비밀이나 응답후의 불이익을 염려하여 거짓 응답하는 것을 조금이나 방지하여 정보의 질을 높이고자 하는 것이다. 앞의 두 회사의 통계 결과에서처럼 거짓된 정보에 의하여 도출된 결과는 신뢰도가 떨어진다.

한편, 본 연구에서 제안하는 방법과 추정 점수는 실제 공기업 등의 청렴도 조사에서 직접적으로 이용할 수 있어 민감한 사항에 대한 모수 값 추정의 효율성을 높일 수 있다.

참고문헌

[1] 박희창, 남기성, 이기성 (2001). Implementation of Randomized Response Technique at Internet Survey, *한국통계학회 논문집*, 8권 3호, pp. 731-738.

[2] 부패방지위원회 (2003). *부패방지백서*, 부패방지위원회.

[3] 류제복, 이계오, 이기성 (1994). 확률화응답기법의 실용화 방안, *응용통계연구*, 8(1), pp. 9-26.

[4] Fox, J. A. and Tracy, P. E. (1986). *Randomized Response : A Method for Sensitive Survey*, Sage Publications.

[5] 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, pp. 520-539.

[6] Greenberg, B. G., Kubler, R. R., Abernathy, J. R. and Horvitz, D. G. (1971). Applications of the RR Technique in Obtaining Quantitative Data, *Journal of the American Statistical Association*, 66, pp. 243-250.

[7] Nam, K. S. Lee, G. S. etc. (2002). *A System for Sensitive Characteristics on Internet Survey,* Asia-Pacific Decision Sciences Institute Conference.

[8] Warner, S. L. (1965). Randomized Response ; A Survey Technique for Eliminating Evasive Answer Bias, *Journal of the American Statistical Association,* 60, pp. 63-69.