

통계 조사의 현황과 전망

Current Situation of Statistical Surveys in Korea and Its Prospects

이 계 오*
Lee, Kay-O

정보화와 지방화 시대에서 통계 이용 실태와 문제점을 요약하였고, 컴퓨터를 이용한 통계 조사의 발달 과정을 살펴보았다. 앞으로 도래할 인터넷 시대에서 전화 조사에 대체될 인터넷 조사의 특징과 발전 방안을 알아보았고, 지방화 시대에서 필요한 지역 통계의 생산 기법으로 소지역 추정법을 소개하였으며, 다양한 학문 분야에서 통계 활용 촉진과 혼란을 방지하기 위해서 통계의 표준화 방안을 제시하였다.

The present status of utilizing statistics and problems in the era of information and localization are summarized and the developing process in a computerized survey is investigated. Also, characteristics and development plans of the internet surveys to come, which can substitute for telephone surveys, are discussed. We introduced small area estimation for producing the local statistics and suggested the standardization of statistics to improve statistics use and avoid misunderstanding in various areas of studies.

I. 서언

미국의 미래학자 Alvin Toffler는 사회의 발달 과정을 농경 사회, 산업 사회 및 정보화 사회로 구분하고 각 과정별 특징을 이렇게 말하고 있다. "농경 사회는 개인의 생존 추구가 목적이었고, 산업 사회는 국가 간 이익 추구를 위하여 과대한 경쟁을 치렀으며 정보화 사회는 공존과 협력을

* 공군사관학교 전산통계학과(kayolee@orgio.net)

추구하기 위해서 정보의 획득과 공유가 필수적인 사회가 될 것이다.” 즉 정보화 사회에서는 신속 정확한 정보의 획득과 활용 능력이 핵심적인 지식임을 언급하고 있다.

21세기의 정보화 사회에서는 정보가 큰 자본이며 지식임을 많은 사람들이 강조하고 있으나 아쉽게도 정보의 인프라스트럭처(Infrastructure)라 할 수 있는 통계의 중요성에 대해서는 별로 언급이 없다. 이는 통계의 기능이나 역할에 대한 많은 사람들의 잘못된 인식이나, 통계에 대한 식상한 인상을 대변하는 것일 수도 있다. 통계 선진국에서도 통계에 대한 잘못된 인식과 오용에 대해서 경고하는 말이 적지 않다. 그 중에서도 Disraeli가 말한 “There are three kinds of lies--lies, damned lies, and statistics.”는 우리가 현 상황에서 되새겨 보아야 할 구절이다.

다양한 분야에서 무절제한 통계 조사와 검증되지 않은 자극적인 통계 숫자들의 범람은 통계에 대한 혐오심을 심어줄 우려가 있기 때문이다. 대부분의 학문 분야에서 통계 활용의 중요성을 인식하고 계량적 분석과 예측적인 연구에는 통계 조사와 분석이 과도하게 강조되는 인상마저 있다. 통계학 분야의 의식있는 사람들은 우려감과 함께 환영하는 태도를 보이고 있다. 우려하는 측면은 통계가 꼭 필요한 분야에서는 미흡하며 사용해야 할 지방 자치 행정이나 지역 발전 계획에는 적절한 통계를 찾아볼 수 없는 반면에 학문 연구나 선거 여론 조사 또는 사회 의식 조사 등에는 검증되지 않는 통계가 홍수를 이루고 있음을 바로 잡아야 한다는 점을 지적하고 있다. 한편 환영하는 측면은 공산품이나 서비스 분야까지도 고객 만족도 등 통계를 이용한 정보를 신뢰하고 있고, 많은 사람들이 통계의 이용 가치를 충분히 인식하고 있음에 대한 책임이 막중해지고 있다는 점을 강조하고 있다. 정보 지식 기반 사회에서 통계 역할의 중요성을 인식하고 다양한 학문 분야에서 통계의 활용도가 증가되는 시점에서, 통계를 많이 다루는 학문 분야간의 통계 관련 정보와 지식을 공유하고 상호 보완적 발전을 추진하기 위한 조사연구학회의 창립은 여러 가지 의미에서 중요한 출발이라고 생각된다.

이제 조사연구학회의 발전적인 출발과 지식정보사회에서 통계의 성공적인 역할 수행을 위해서 우리나라의 통계조사 현황을 살펴보고 앞으로 발전 방안을 모색해 보는 것도 가치 있는 일이라고 생각되어 본 주제를

발표하고자 한다.

현황에서는 먼저 다양한 분야에서 통계가 활용되고 있는 실태를 알아보며, 다음에는 조사 방법의 장단점을 살펴보고, 발전 방안에서는 정보통신 기술의 발달로 폭발적인 수요가 예상되는 인터넷 조사의 개념, 장단점과 발전 방안을 알아보고, 지방자치 행정에서 필요한 지역 통계의 개발 방안을 언급하며 마지막으로 통계 표준화의 필요성을 제안하였다.

II. 통계 조사의 현황

통계 이용 실태를 살펴보기 위해서 먼저 공식 통계를 생산하고 있는

〈표 1〉 부문별 작성 현황

부문	종수	구성비	조사 통계	보고 통계	분석가공 통계
인구	21	5.3	3	16	2
고용, 임금	21	5.3	20	1	-
물가, 가계 소비	16	4.0	16	-	-
보건, 사회 복지	70	17.5	27	41	2
환경	16	4.0	7	9	-
농림, 수산	53	13.3	28	23	2
광공업, 에너지	18	4.5	11	7	-
건설, 주택, 토지	25	6.3	15	10	-
교통, 정보 통신	41	10.3	9	32	-
도소매, 서비스	6	1.5	6	-	-
경기, 기업경영	44	11.0	40	-	4
국민지역계정	9	2.3	1	1	7
재정, 금융	15	3.8	2	13	-
교육, 문화, 과학	15	3.8	10	5	-
기타	30	7.5	8	4	18
합계	400	100.0	203	162	35

통계청의 자료를 알아보면 1999년 11월 1일 기준으로 통계청에서 관리하는 통계는 400종으로써 정부 기관에서 작성하는 통계는 276종(74%), 지정 기관에서 생산하는 통계는 103종(26%)이다. 이를 작성 방법별로는 조사 통계는 203종(51%), 보고 통계는 162종(40%), 분석 가공 통계는 35종(9%)이다. 이를 부문별로 구분하면 <표 1>과 같다.

국정을 운영하는데 필요한 국가적 단위의 통계는 대부분 명목상으로 생산되고 있지만 산업구조의 변화에 따른 서비스 산업의 비중이 커지고 있으나 서비스 관련 통계는 미흡하고, 사람들이 피부로 느끼는 체감과 공식통계간의 차이는 많은 논란이 되어 왔다. 특히 IMF 이후 실업 문제가 사회적 이슈가 됨에 따라서 실업 통계의 개념과 숫자에 대한 많은 문제점들이 제기되었다.

몇 가지 미비점이나 개선 보완해야 할 사항이 있음에도 국가단위 통계는 통계 중진국 이상의 수준에 있다고 볼 수 있으나 지방자치 제도 실시와 더불어 각 지방 정부에서 필요로 하는 모든 분야의 지역 통계는 양적으로나 질적으로 아직 많은 문제점들을 가지고 있다.

주민들의 선거에 의해서 자치 단체장들이 선출되다 보니 각 지방자치 정부들은 해당 지역 특성에 적합한 장기지역개발계획을 수립해야 하고 또한 지역 주민들에게 필요한 지방행정을 수행해야 하므로 해당 지역을 잘 나타내는 통계가 필요하지만 이를 생산해줄 통계 조직이나 통계전문 인력은 빈약한 실정일 뿐만 아니라 통계는 중앙정부에서 공급해 주는 것으로 잘못 인식하고 있는 경우도 적지 않다. 앞으로 지방자치행정에서 필요한 통계를 생산할 수 있는 방안은 다음절에서 자세하게 언급하겠다.

다음에는 일반 여론조사 기관이나 연구소 등에서 무분별하게 생산하여 공표 되고 있는 잘못된 통계의 활용 실태를 살펴보자. 우선 사회적으로 많은 관심을 받고 있으며 또한 파급 효과가 큰 선거여론 조사는 1987년 12월 17일 대통령 선거에서 각 후보에 대한 지지율과 득표율을 개표 시작 전에 정확하게 예측함으로써 각종 선거에서 많은 각광을 받아 왔고 또한 여론조사 기관의 성장과 발전을 거듭하였으나 새로운 기법이나 이론적인 연구를 등한시 한 결과 1996년 4월 11일 국회의원 선거에서는 영터리 예측 보도로 세계적인 뉴스 거리가 되었다. 이는 여론조사의 실패

사례로서 많은 책이나 연구 논문에서 인용되기도 하였다. 엉터리 예측 보도를 하게된 원인은 여러 가지가 있겠으나 대표적인 것은 여론조사 기관과 방송사 관계자들의 통계에 대한 인식 부족과 과도한 경쟁 방송으로 볼 수 있다. 여기에서 교훈은 통계조사를 담당하는 사람이나 이용하는 사람 모두에게 통계에 대한 기초적인 지식과 윤리 의식이 필요하다는 것이다.

우리 통계수준을 나타내는 잘못된 통계의 한 사례는 국내의 인터넷 사용 인구수에 대한 보도이다. 얼마전 모 일간지에 보도된 내용 중에서 많은 사람들의 관심의 대상이 된 인터넷 사용자 수를 보면 정보통신부는 '98년 말 기준으로 310만 여명에서 99년 9월에는 520만 명으로 증가했다고 발표했으나 갤럽과 온 디지털은 13세에서 59세 사이의 총인구의 26.1%에 달하는 820만 명으로 추정되었다고 발표하였고 현대 리서치 연구소는 420만 명이 인터넷을 사용하고 있다고 추정하였다. 어느 통계가 맞는 것일까? 오차가 무려 100%에 달하는 통계를 어떻게 신뢰할 수 있을까? 모든 사람들이 인터넷 사용 인구는 폭증하고 있다는 데는 동감하지만 인터넷 사용자의 추정수가 조사기관에 따라 420만 명에서 820만 명까지 차이가 난다면 앞으로 인터넷을 이용한 통계 조사가 보편화 될 경우에는 더 큰 비표본오차가 발생할 것이다. 따라서 통계 작성과 발표 과정을 표준화하지 않는다면 앞으로 통계 이용에 많은 어려움이 예상된다.

금년 10월부터 시작된 국민 인터넷 PC 보급정책의 실시로 서민 계층에 100만 원대 인터넷 PC가 보급되고 인터넷 사용료가 점차 저렴해지면 인터넷 사용 인구는 급증하여 2000년 말경에는 1000만 명을 넘어설 것으로 예상되고 이에 따라 인터넷을 이용한 조사영역은 점차 확대될 것이다. 인터넷 조사 방법을 언급하기 전에 컴퓨터를 이용한 조사 방법을 알아보는 것이 인터넷 조사 방법의 장단점을 이해하는데 도움이 될 것이다.

컴퓨터 과학 기술의 발달과 더불어 컴퓨터보조 조사정보 수집(CASIC : Computer Assisted Survey Information Collection) 방법도 빠른 속도로 발달하였다. CASIC의 최초 시도는 1970년 초에 컴퓨터보조 전화면접(CATI : Computer Assisted Telephone Interviewing)이 대규모 메인 프레임을 갖

춘 조사기관에서 실시되었으며 이때 보고된 내용은 연필과 종이를 사용한 전화조사와 비교해서 비용 절감, 자료의 질적 향상 및 적시성에서 진전되었다는 것이었다.

컴퓨터 기술의 발달로 컴퓨터 가격이 대폭 낮아지고 소프트웨어도 낮은 가격으로 공급되면서 대학연구소와 여론조사기관들도 CATI에 의한 마케팅 조사의 비중이 커졌다. 1990년 초까지 여론조사나 마케팅 조사들 중에서 90% 정도가 전화 조사를 실시하고 있었고 전화 조사중에서 CATI에 의존하는 비율이 60%이상이며 우리나라에서도 갤럽 등 여러 조사기관들이 CATI를 도입하고 있으나 많은 경우에는 CATI와 종이와 연필을 이용한 전화조사를 병행하여 실시하고 있다.

면접조사에서도 컴퓨터를 이용함으로써 면접원으로 인한 오차를 줄이려는 시도가 1980년 초부터 미국 등에서 시작되었다. 컴퓨터보조 개인면접(CAPI : Computer Assisted Personal Interviewing)은 컴퓨터의 무게가 가벼워지면서 면접 프로그램이 내장된 노트북이나 휴대용 컴퓨터를 휴대하고 면접 대상자를 방문하여 개인적으로 민감한 사항 등에 대해서도 응답자들이 컴퓨터에서 직접 설문을 읽고 응답하는 조사를 실시하였다.

1990년 대 초에 CAPI의 보급은 급속도로 확산되었다. 미국에서 노동력 조사 등 정부통계 조사에 활용되면서 세계적으로 확대 적용되었으며, 우리나라에서는 작년부터 경제활동 인구조사에서 노트북을 이용한 CAPI 형식의 조사 방법을 시행하고 있다. CAPI 활용은 조사비용의 절감과 조사자료의 입력 과정을 생략하는데서 오는 시간 절약과 자료입력과정에서 발생 가능한 오차를 줄일 수 있는 장점도 있어 앞으로 인터넷 조사와 병행하여 실시함으로써 범위 오차(Coverage Error)를 보완할 수 있는 조사방법 중 하나가 될 것이다.

응답자의 개인적 비밀 사항이나 민감한 사항에 대한 조사에서 좀 더 효과적으로 시행할 수 있는 방법이 컴퓨터자제식 자료수집(CSAQ : Computerized Self- Administered Data Collection)이다. 이는 면접 과정을 디스켓에 저장하여 조사대상자에게 전달하면 조사대상자는 이를 자신의 PC에 설치한 후 프로그램에 따라 응답한 결과를 디스켓으로 조사기관에 우송하거나 전자 메일을 이용하여 조사기관의 컴퓨터로 보내는 방식이

다. 물론 응답자가 자신의 컴퓨터를 보유하고 있어야 하며 면접조사 프로그램을 설치하여 응답할 수 있어야 한다는 제한 사항이 있지만 컴퓨터의 보급이 대중화되고 컴퓨터 사용 능력이 향상되어 가는 추세임을 감안할 때 CSAQ는 여러 가지 형태로 발전할 것으로 예상된다. CAPI의 발전된 형식으로 컴퓨터보조 자기면접(CASI : Computer Assisted Self-Interviewing)은 응답자의 대답 내용에 따라서 질문의 순서를 통제하거나 응답 내용을 컴퓨터가 점검하여 자료의 질을 향상시킬 수 있는 방법이다. 그밖에 무료전화번호를 이용하여 컴퓨터가 설문을 읽어 주면 응답자가 듣고서 전화기의 번호를 눌러서 응답함으로써 응답자와 컴퓨터간의 반복적인 질문-응답이 가능하고 개인적으로 민감한 사항을 조사할 수 있게 하는 전화자료입력(TDE : Touchtone Data Entry)방법과 자료 입력을 전화기 다이얼 대신 음성으로 응답하는 음성인식입력(VRE : Voice Recognition Entry) 방법이 있다. TDE 방식은 우리나라에서도 여론 조사 기관이나 마케팅 조사기관에서 활용하고 있는 방법이다.

위에서 언급하지 않는 일반적인 우편조사와 전화조사 등은 우리나라의 대학 연구소, 소규모의 조사기관 또는 연구기관 등에서 많이 이용되고 있다. 우편조사와 전화조사는 많은 장점도 있으나 취약한 점은 응답률이 너무 낮다는 것이다.

Ⅲ. 발전 방안

통계조사 현황에서 지방자치제의 정착에 발맞춘 지역통계의 활성화 및 정보화 사회에서 정보 홍수 속에서 무차별하게 남발되는 통계의 표준화 필요성 그리고 컴퓨터의 발달과 인터넷의 확산으로 새로운 조사방법으로 등장한 인터넷 조사의 표준 절차에 대한 연구의 필요성에 대하여 알아보았다. 본 절에서 인터넷 조사의 개관, 조사방법과 표준 절차를 먼저 다루어 보고 다음으로 지역통계 활성화 방안과 조사통계의 표준화 방안에 대해서 간략하게 설명하겠다.

1. 인터넷 조사

급증하는 인터넷 인구를 대상으로 인터넷망을 이용하여 자료를 수집하는 조사방법을 인터넷 조사로 총칭하지만 기존 조사방법에 대해 큰 차이점은 조사 대상이 국내 또는 특정계층으로 제한되지 않고 국경이나 공간의 제한을 받지 않는다는 점이다.

인터넷 조사는 전자 메일의 이용에서부터 시작되었다. 1980년 이후부터 소수의 특정 집단이기는 하지만 전자 메일을 통한 간단한 설문조사를 시행하기 시작하였으며 네티즌의 인구가 급증한 1990년대에 들어와서 보편화되기 시작하였고 특히 WWW(World Wide Web) 사용이 확대되면서 여러 가지 형태의 인터넷 조사방법이 연구되기 시작하였다.

인터넷 조사방법과 특징을 먼저 살펴보고 인터넷 조사의 표준 절차를 설명하겠다.

1) 인터넷 조사의 특징

컴퓨터와 정보통신 공학의 발달로 인터넷 사용의 비용이 저렴해지고 사용에 익숙해지면서 교육 수준과 생활 수준측면에서 폭넓은 대상으로 확대되고 있다. 네티즌의 직업도 컴퓨터 관련 직종에서 다양한 직업으로 확대되고 남성 편중 현상이 갈수록 낮아지면서 성별이나 연령대별 불균형적 분포가 완화되어가고 있다. 인터넷 조사의 장점은 아래와 같이 요약할 수 있다.

- ① 비용이 갈수록 저렴해지고 있으며 대략적으로 전화조사의 50% 정도이고 DBM (Disk-by-Mail)을 이용한 우편 조사의 20% 정도가 소요된다.
- ② 조사소요기간도 기상상태의 영향을 받지 않을 뿐 만 아니라 대폭 단축되고 있으며 DBM에 의한 우편조사는 4-6주일이 소요되고 전화조사는 2-3주일이 소요되지만 인터넷 조사는 2-3일이면 충분할 것이다.
- ③ 확보된 Frame에서 표본 추출이 용이할 뿐 만 아니라 설문 문항의 표현이 다양하고 응답자가 이해하기 쉽도록 그래픽이나 음성 또는

동화상을 이용한 설문지 설계가 가능하다.

- ④ 응답자와 대화 형식을 통해서 정성적인 자료도 용이하게 수집할 수 있고 응답자의 반응에 따라 되묻기를 할 수도 있다.

인터넷 조사방법의 개선과 컴퓨터 기술 및 소프트웨어의 발달로 더 많은 장점들이 있을 수 있으나 현재의 수준에서 몇 가지의 장점을 알아보았으며 반면에 인터넷 조사 이용에서 유념해야할 단점도 아래와 같다.

- ① 응답자를 통해서 관찰되지 않은 자료 수집은 불가능하다. 무응답에 대한 반복적인 설문지 전송은 할 수 있으나 방문이나 기타 방법을 통한 자료 수집이 곤란하다.
- ② Web 페이지를 이용한 조사에서는 간단한 내용의 질문만 가능하므로 조사 도구에 따라서 많은 제한이 있을 수 있다. CSAQ 관련 인터넷페이지, 컴퓨터 운영체제나 사용 브라우저에 따라서 호환성에 제한을 받을 수도 있다.
- ③ 일반적인 모집단을 대표할 수 있는 표본 추출틀을 얻을 수 없는 경우가 많다. 좀더 자세한 내용은 뒤에서 언급하겠다.

인터넷 조사의 계획 시 적합성 검토에서 최소한 다음 3가지 사항은 점검해야 한다.

첫째 모집단을 대표하는 표본 추출이 가능한가?

현재 상황에서 대부분의 경우에는 편향된 표본 추출이 우려된다. 만일에 편향된 표본을 추출하였다면 이에 대한 보완 수정할 수 있는 추정법을 적용해야한다.

둘째 Web에 사용할 질문과 자료들이 비밀 사항은 아닌가?

만일에 비밀스러운 내용이 있다면 중앙 통제 방식을 채택함이 좋을 것이다.

셋째 조사 과정이 복잡하고 장시간이 소요되지 않는가?

질문 과정이 15분 이상 소요될 때는 무응답과 오답 확률이 높다는 점에 유의해야 한다.

인터넷 조사에서 정확한 응답을 얻기 위해서는 질문 형식을 잘 선택해야 한다. 질문 형식에 따라서 자료처리 및 분석의 용이성과 비표본 오차의 감소 정도가 다르다. 예를 들어 2지 택일이나 4지 택일과 척도화된 질문과 같이 단일 응답을 요하는 질문형식은 자료 처리와 오차 감소 차원에서 인터넷 조사에 제일 적합한 것이다. 반면에 단답 서술형이나 설명형과 같은 질문 형식을 인터넷 조사에서 사용할 경우에는 무응답률이 높아질 수 있음을 명심해야 한다.

2) 인터넷 조사 방법

네티즌을 대상으로 한 조사방법은 관점에 따라서 여러 가지로 분류할 수 있으나 본고에서는 대화형 프로그래밍(Interactive Programming)의 정도에 따라서 3가지로 구분하여 설명하겠다.

첫째는 플랫폼 파일(Flat File)형식으로 전자메일조사를 포함하여 현재 각 신문사와 인터넷 회사들이 활용하는 방법으로 Web 페이지에 간단한 설문을 게재하고 Web 페이지를 방문한 사람들을 대상으로 의견을 청취하는 것이다. 여기에서는 대화형 프로그래밍을 일체포함하지 않기 때문에 프로그램 유지비용이 적고 대부분의 네티즌들이 익숙한 형식이다. 좀 더 복잡한 내용을 묻거나 질문의 유형과 문항수의 제한이 크다는 단점이 있다.

두 번째는 대화 형식으로 플랫폼 파일형식과 유사하지만 서버 사용자와 대화형 프로그래밍이 연결된 대상자들에게 사용하는 방법으로 앞에서 응답한 질문과 앞부분의 질문을 검색하여 질문 순서를 바꾸어가면서 응답할 수 있는 것이 특징이다. 인터넷 조직이 확충되고 CAPI 소프트웨어를 장착한 서버에 연결된 경우에 유익하고 인터넷의 속도와 자료의 전송에서 신뢰성과 보안성이 보완 발전된다면 제일 많이 활용될 수 있는 형식이다.

미국에서 취업통계조사(CES : Current Employment Statistics)에서 적용하고 있는 방식도 여기에 속한다고 말할 수 있으며 CES에서는 표본 기업체에게 팩스 또는우편엽서로 조사 일정을 통보하고 표본기업체는 조사센터의 Web 페이지에 접속하여 설문 문항 내용에 응답을 마친 후

“Submit”을 클릭(Click)하면 자료들은 조사센터의 컴퓨터에 자동 입력되고 입력된 자료의 논리성 여부를 검사하여 보완 수정해야 할 조사기업체와 무응답 기업체에게는 전자 메일 또는 팩스/우편엽서로 재차 조사에 응할 것을 알림으로써 응답률을 높이고 있다.

세 번째는 온라인 대화(On-line Chat)형식으로 동일컴퓨터 조직에 연결되어 개인간의 메시지를 상호 교환하면서 대화 형식으로 조사를 실시할 수 있는 방법이다. 현재는 제한적으로 사용되고 있으나 인터넷 속도가 빨라지고 비디오 이용이 가능해진다면 정성적 특성 연구에 많이 이용될 전망이다.

3) 인터넷 조사설계 원칙

인터넷 조사에 전자메일조사도 포함되고 있으나 본 절에서는 Web 페이지를 이용한 인터넷 조사에 초점을 맞추어 Web 조사설계에서 지켜야 할 원칙들을 알아보겠다. Web 조사의 설계는 모든 조사 대상자에게 설문지가 잘 전달되고 설문지를 접한 응답자들은 혼쾌하게 응답할 마음이 생겨서 정확하게 응답하는데 도움을 주도록 되어야 하고 응답자들의 컴퓨터 사용 기술수준이나 연구 주제에 대한 관심의 유무에 영향을 받지 않는 조사설계를 이상적인 것으로 간주했을 때 Web 설문지 작성 원칙은 다음과 같다.

- ① 감사 인사로 시작하고 응답 요령이 쉽다는 것을 강조하며 페이지를 넘기는 요령을 자세하게 설명하라.
- ② 설문의 첫 페이지는 모든 응답자들이 쉽게 이해하고 간단하게 응답할 수 있는 질문을 한 화면에 채우도록 하라.
- ③ 설문에서 문항 배열 요령이나 질문의 형식은 통상적인 종이설문형식과 유사하게 작성하라.
- ④ 장황한 설명으로 문장이 길어서 브라우저 화면에서 겹치지 않도록 문장의 길이를 줄여라. 문장이 길면 읽을 때 건너뛰는 경우가 있을 수 있으나 짧으면 건너뛰는 경우가 없다.
- ⑤ 설문의 응답에 필요한 컴퓨터 조작 방법은 별도 지시로 설명하라. (문단을 달리하거나 글자체를 달리하여 주의를 환기 시켜야 한다.)

- ⑥ 각 문항별로 응답에 필요한 컴퓨터 조작 요령을 설명해야 하며 질문지 앞부분에 묶어서 설명하는 것을 피하라.
- ⑦ 모든 질문에 대해서 핵심적인 답을 요구하는 질문 형식은 삼가고 응답 항목으로 “기타” 또는 “잘 모름”이라는 항목을 포함시켜라.
- ⑧ 질문의 순서가 중요한 의미가 없다면 문항별로 스크롤(Scroll)할 수 있도록 설계하고 필요한 경우에는 전화조사 등과 혼합하여 사용할 수 있도록 설문을 설계하고 자료는 결합해서 분석토록 해야 한다.
- ⑨ 선택 문항수가 많아서 한 화면에 나타낼 수 없을 때는 적절한 설명을 붙여서 더블 बैं킹(Double -banking) 또는 트리플 बैं킹(Triple-banking)을 고려하라.
- ⑩ 설문이 거의 끝나는 상황일 때 이를 응답자에게 알려 주는 말이나 기호를 사용하여 중도에서 응답을 포기하지 않도록 하라.(만일 2-3개 문항이 남아 있는 상태에서 응답을 중도 포기한다면 무응답이 많아 질 것이다.)
- ⑪ 적절하다고 생각되는 항목을 모두 표시하시오 또는 개방형 질문을 사용할 때 응답 오차가 발생할 수 있음에 유의하라.

위에서 언급한 원칙들은 응답자들 중에는 컴퓨터 장비의 낙후, 전송 환경의 미흡과 브라우저의 문제로 설문지를 받지 못하거나 응답할 수 없는 경우를 고려하였고, 설문에 응답하는 요령과 컴퓨터 조작 절차의 숙련정도는 응답자의 입장을 고려하였으며 필요한 경우에는 우편조사 등과 혼합한 조사방법을 적용할 수 있음에 유의하여 작성하였다.

2. 지역 통계 활성화

정보화와 지방화라는 정치 사회적 발전 과제를 해결하는데 조사연구 학회의 역할이 기대되고 또한 인터넷 등 정보통신수단의 발달로 세계의 지구촌화가 빠르게 진행되고 있으므로 지역 간의 균형적 발전을 위해서 지방자치 행정이 성공적이고 안정적으로 수행되어야 할 것이다. 지방자치 행정의 효과적인 수행을 위한 지역통계의 활성화가 우선적으로 해결

되어야 할 것이다.

1) 지역 통계의 확충

지역의 특성과 현황을 파악하기 위한 통계를 개발하고 작성하기 위해서 지역 통계 생산 조직이 보강되어야 하지만 이는 단기간에 이루어질 수 있는 사안이 아니다. 현재 여건 내에서 추진할 수 있는 일들은 지방 행정 업무에서 파생되는 자료를 통계화하고 각 지역 내에 있는 상공회의소나 한국은행 지점 등의 지정통계기관을 활용하여 지역경제 부문에서 필요한 통계를 작성해야 한다. 다만 지방 정부에서 필요한 통계를 공유하고 국가 단위의 통계는 각 시도의 통계를 단순한 합산으로 만들 수 있도록 통계에 대한 인식과 제도를 시급하게 정비해야 한다.

각 시도별로 지방정부에 필요한 통계의 종류를 정확하게 파악하고 이를 효과적으로 생산할 수 있도록 통계유관기관(예를 들어 지방 대학의 통계 관련 학과, 각 지방 발전 연구원, 행정 조직 및 통계 지정 기관의 지방 사무소 등)이 함께 참여하는 지역통계 발전협의회(가칭) 등의 조직을 편성 운영한다면 각 시도별로 학·연·관 간의 긴밀한 협력 활동을 보다 효과적으로 추진할 수 있을 것이다.

통계청이나 행정기관들이 주관하는 전국 규모의 통계조사는 각 지방 정부가 위임받아서 수집한 자료와 각 지방 사무소에서 실시한 해당 지역의 통계조사를 참조하고 있으나, 각 시도의 지방정부에서 필요한 조사항목을 추가하여 신뢰성과 시의성을 갖춘 지역통계를 생산할 수 있도록 조사체계를 개선 보완하고 각 지역별로 표본을 층화해서 지역통계와 국가 단위의 통계를 함께 생산할 수 있도록 해야 할 것이다. 특정 지역에서만 필요한 통계항목은 해당 지방 정부가 독자적으로 자체 통계조사 활동을 계획하여 지역통계를 생산해야 하고 이때도 지역통계 협의회의 자문과 심의를 거친다면 신뢰성을 갖춘 지역통계를 생산할 수 있을 것이다.

2) 지역 통계의 질적 개선

각 지방정부의 독자적 통계조사와 행정 업무자료의 공식적인 지역 통계화를 위해서는 용어, 개념, 기준, 통계 항목, 조사 방법, 조사 대상 등

에 관한 기준과 지침서를 작성하여 지역 통계 작성의 절차를 규범화함으로써 공식통계로서 필요한 최소한 질적수준을 확보해야 할 것이다. 유럽 국가들의 많은 공식통계는 조사활동을 통해서 얻어지기보다는 업무자료를 표준화한 보고통계로 구성되어 있다. 특히 정보화 시대에 필요한 지역통계의 데이터 베이스를 구축하기 위해서 분류체계 및 편집체계를 표준화하여 코드화가 용이하도록 해야 한다.

각 시도에서 지역통계 생산에 필요한 통계전문 인력을 별정직 공무원으로 채용하여 통계전담 전문인력으로 활용해야 하고 만일 불가능한 경우에는 각 지방 대학간 학·관의 협력 체계를 구축하여 석사·박사 과정의 전문 인력을 활용한다면 실용적인 통계교육과 연구에도 도움이 되고 지방 행정을 효율적으로 수행할 수 있으므로 경제적으로도 도움이 될 것이다. 장기적으로 각 지방정부에 통계 업무를 담당하는 부서를 통계업무 전담별정직으로 관리할 수 있는 인사관리를 추진해야 할 것이다. 각 지방 정부의 경쟁력을 제고하기 위해서, 전국을 하나의 정보시스템으로 연결하는 지역정보 시스템 구축을 위한 지역통계 DB를 준비해야 한다.

각 지방정부 별로 지역통계 DB를 개발하게 되면 많은 예산과 전문인력이 필요하게 되므로 열악한 지방 정부의 재정으로는 감당할 수 없기 때문에 중앙정부에서 예산과 전문인력을 지원하고 시범적으로 특정 시도를 선정하여 지역통계 DB 구축사업을 추진한 후에 이를 표준으로 전국적으로 확대 운영할 수 있도록 해야만 각 지방정부 DB간의 호환성을 제고할 수 있을 것이다.

3) 지역 통계 작성 기법 개발

지방자치단체에서 유용하게 활용할 수 있고 최소 비용으로 생산이 가능한 지역통계 작성기법을 개발하여 지역통계의 활용성을 높일 수 있는 방안을 알아보자. 새로운 지역통계기법 개발 연구는 지방자치단체에서 주관하기에는 예산이나 연구 능력 등을 고려할 때 어려움이 많을 것이므로 중앙 통계기관인 통계청에서 적극적으로 지원해야할 사항이다. 여기서 지역 경제의 영향을 일목요연하게 파악할 수 있는 산업 연관표와 지역통계 작성에 활용될 수 있는 소지역 통계추정에 관해서 언급하고자 한다.

① 지역산업연관표 작성

정보화 사회가 심화될수록 경제적 환경의 변화는 다양해 질 것이다. 지역 간의 변화에 따른 영향과 지역내 산업간의 변화의 영향은 상당히 높은 상관성을 갖고 있기 때문에 이들을 종합적으로 평가 분석하여 파악하지 못한다면 효과적인 정책결정에 어려움이 따를 것이다. 특히 한 지역에서 국제적 행사나 신규산업진흥 추진계획을 수립할 경우에는 주변 지역 간이나 산업간의 파급효과를 계량적으로 분석할 수 있는 산업연관표를 작성해야 할 것이다.

일반적으로 산업연관표는 생산과정에서 원재료 등에서 중간 소비된 것까지 포함되도록 일정기간 동안 지역내의 경제 활동에서 발생하는 재화와 서비스의 생산·처분과 관련된 모든 거래 내역을 일정한 원칙과 형식에 따라 기록한 통계표로서 그 지역의 경제 흐름과 구조를 자세히 나타내는 일종의 해부도적인 성격을 갖는다. 이러한 산업연관표는 현재 한국 은행에서 국가 단위로 작성하고 있으나 적어도 각 시·도 단위까지 작성되어야 각 시·도에서 지역정책 개발계획의 참고자료로 활용할 수 있을 것이다. 산업연관표와 유사한 특성을 갖은 지역연관표가 작성되어야만 지역의 투자계획을 세우거나 지역 간의 물류 운송문제를 분석할 수 있으며 이들 연관표를 이용하면 지역 간 교역되는 재화 및 서비스의 지역 간·산업간 거래의 경제효과를 정확하게 평가할 수 있을 뿐만 아니라 지역별 산업 연관 분석이 가능하다. 지방자치제의 선진국인 일본에서도·부·현별 산업연관표를 작성하여 지역경제 정책결정에 활용하고 있다.

② 소지역 통계 추정(Small Area Estimation)

우리나라의 통계는 그 동안 총량적인 성장위주의 국가발전 전략에 초점을 맞추어 주로 중앙정부 주도의 국가단위 통계를 중심으로 발전시켜 온 덕분에 국가단위 통계는 선진국 수준까지 접근하였다. 전국단위 또는 시도단위 통계를 위한 대규모 통계조사가 정기적으로 실시되고 있으므로 이를 이용한 지역통계를 생산할 수 있는 방법으로 소지역 통계 추정법에 대해서 간략하게 소개 하고자 한다.

소지역 통계 추정 방법은 미국이나 캐나다에서 현재 적용되고 있는 지역통계 생산 기법으로 지금 통계학계에서는 연구가 활발하게 진행되고 있다. 먼저 소지역 통계추정의 개념과 추정방법의 종류를 살펴보자.

소지역 통계 추정이란 모집단의 모수추정을 위한 통계조사를 실시한 후에 모집단보다 좁은 소지역의 특성을 추정하는데 사용되는 기법이며 여기서 소지역이라 함은 모집단에 비해서 상대적으로 좁은 영역을 갖는다는 의미보다는 관심대상의 영역에 속하는 표본의 규모가 적음을 뜻하기 때문에 규모가 적은 표본에서 특성치를 추정한다면 분산이 커지므로 신뢰할 만한 추정값을 얻을 수 없을 것이다. 소지역 추정방법은 추정치를 산출하려는 영역과 유사한 다른 영역으로부터 정보를 빌려서 관심 모수에 대한 추정치의 정도를 높이는 것은 물론이고 해당 소지역의 보조정보와 결합하여 추정값의 정도를 높이는 특징도 있다. 소지역 추정방법은 크게 두 가지로 분류할 수 있다. 첫째는 해당 소지역에서 관찰된 표본자료들을 이용하여 추정값을 구하는 직접 추정법(DE : Direct Estimation)이며 두 번째는 해당 소지역에서 추출된 표본의 관찰자료와 이와 유사한 다른 소지역의 관찰자료와 보조정보를 이용하여 추정량을 구하는 간접 추정법(IE : Indirect Estimation)이 있다.

미국의 경우에는 인구 센서스 자료를 이용하여 카운티(County)당 추계 인구수의 추정과 인공위성 사진과 표본조사의 자료를 이용한 카운티당 옥수수과 콩의 재배면적 추정 등에 소지역 통계추정 방법을 적용하고 있다. 우리나라는 경제활동 인구조사(경활)를 통해서 시·도 단위까지 취업, 실업, 노동력 등과 같은 인구 경제적 특성을 조사하고 거시 경제 분석과 인력자원의 개발정책 수립에 필요한 기초자료인 고용구조, 가용노동시간 및 인력자원 활용도를 추정하여 고용창출, 취업훈련 소득증진 등의 정책입안 및 평가자료를 활용하고 있다. 경활자료와 노동력 관련 DB 자료를 이용하여 시·군·구 단위의 실업자와 산업 경제 관련 지역통계를 생산하는데 소지역 통계 추정방법을 적용할 수 있는 타당성을 검토하여 다양한 분야의 지역통계생산에 적용 가능한 소지역 추정법의 연구가 필요하다.

3. 통계의 표준화

통계 조사와 관련된 여러 분야에서 학문을 익힌 사람들이 한 자리에 모였다. 물론 정확하고 신뢰성있는 통계를 올바르게 활용함으로써 사회 발전에 기여할 뿐 만 아니라 올바른 길로 선도하자는 뜻을 가지고 조사 연구학회를 창립하였다. 본 학회가 해야 할 중요한 일이라고 생각되는 것 중 하나는 통계의 표준화이다. 물론 통계 표준화는 간단하게 설명될 수 있는 사안은 아닐 것이나 본인은 두 가지 내용을 제안하고자 한다.

첫째는 통계 용어의 통일 또는 표준화이다. 통계 조사에 관련된 책과 논문들을 읽어보면 사람에 따라 또는 학문에 따라 서로 다른 용어를 사용하고 있다. 정보통신의 발달로 국경과 공간의 한계 없이 정보와 지식들은 공유하고 있다. 그러나 통계 조사에 관련된 많은 대학 교재에서 동일한 의미가 서로 다른 용어로 설명되므로 다른 책을 참고하는데 어려움이 적지 않을 것이다. 예를 들면 샘플링(Sampling)이 “표집”, “표본 추출”, 또는 “샘플링”이라는 용어로 사용되는가하면, 랜덤 샘플링(Random Sampling)은 “무선 표집”, “임의 추출”, 또는 “랜덤 추출”이라고 사용되는 경우를 우리는 쉽게 발견할 수 있다. 이는 하나의 예에 불과하지만 대학에서 공부하는 학생과 일상 사회생활 속에서 통계를 접하는 국민들의 통계에 대한 이해와 이용을 활성화하기 위해서는 통일된 용어 사용이 필수적이라 생각된다. 다음에는 통계조사 결과의 보도 또는 연구논문에서 인용할 때 표준화된 형식을 갖추어야 할 것이다. 물론 언론이나 방송은 상업성이 목적이므로 일반 대중들의 관심과 시선을 끌기 위해서 자극적이고 극단적인 표현이나 내용을 핵심적으로 다루고자 할 것이다. 그러나 이는 통계에 대한 불신을 조성하게 되어 건전하고 신뢰성 있는 통계의 발전을 막는 요인이 될 수 있음을 알아야 한다.

여론조사 등에서는 이런 피해를 줄이기 위해서 유럽 마케팅 여론조사협회(ESOMAR : European Society of Opinion and Marketing Research)에서는 조사결과를 발표하는데 지켜야 할 규칙들을 규약으로 정하기도 하였다.

예를 들면 앞으로 언론이나 방송에 보도되기 위해서는 통계 조사에 의한 결과들은 최소한 모집단(조사 대상), 표본 크기와 구체적인 조사 지

역, 조사 시기와 조사 방법에 대해서는 공개적으로 언급해야 한다 라는 규약이나 지침을 제정하여야 할 것이다.

또 한 가지 통계의 표준화에 도움을 줄 수 있는 방안은 앞으로 내년 3월경에 시행될 예정인 “사회 조사 분석사 II급” 자격 제도를 활용하여 어느 일정 규모의 조사 기관이나 전국 규모의 언론이나 방송 등에서 보도할 수 있는 통계 조사의 결과들은 일정한 자격을 갖춘 사람에 의해서 인증을 받도록 하는 방법도 생각해 볼 수 있을 것이다.

몇 년 전에 본인이 정부 기관이나 공공기관에서 발표한 400여 종의 통계를 분석한 결과에 의하면 비표본 오차 관리 측면에서 볼 때 10종 미만의 통계만이 일정한 발표 형식(조사 목적 명시, 조사 대상과 기간, 조사 방법, 모집단과 표본 프레임, 표본 추출 방법과 추정 오차의 명시)을 갖추었을 뿐 대다수의 통계는 재인용이나 분석 가공 통계를 생산하는데 전혀 이용될 수 없음을 알 수 있었다. 앞으로 통계의 공용화와 DB화를 위해서는 표준화된 통계의 생산을 위한 제도적인 규약이나 지침은 필수적이고 시급한 과제일 것이다.

IV. 결론

조사연구학회의 창립을 계기로 통계조사의 현황과 앞으로 해결해야 할 과제와 요즈음 폭발적으로 증가하는 네티즌을 대상으로 한 인터넷 조사에 대해서 살펴보았다.

현황에서는 통계가 정보화 사회에서 인프라 스트럭처의 기능을 제대로 못하는 실태를 분석하였고 정치·사회의 발전에 맞추어 필요한 지역 통계의 확충의 필요성을 문제로 제기하였으며 개발의 필요성과 컴퓨터를 이용한 조사 방법의 발달 과정 및 특성을 요약하였다.

발전 방안에서는 요즈음의 화두인 인터넷 조사에 대한 개념과 특징을 요약하고 인터넷 조사 방법을 3가지로 구분하여 장단점과 발전 전망을 살펴본 후에 올바른 인터넷 조사를 위한 Web조사 설계의 원칙을 설명하였다. 지방자치화 시대에서 경제적이며 효과적인 지역 통계 작성 기법인

소지역 추정법을 중심으로 지역 통계 활성화 방안을 제시하였으며 통계의 신뢰성과 대중성을 제고하는 방편으로 통계의 표준화라는 과제를 제기하였다.

다양한 학문적 배경을 가진 사람들이 통계라는 매개체를 통해서 통계 조사의 발전을 위한 협동적 노력을 다할 수 있는 조사연구학회가 되었으면 하는 마음으로 문제 제기와 해결 방안을 제시하였지만 이는 순수한 개인적인 견해이므로 혹시 오해나 잘못된 점이 있다면 상호간의 대화 또는 의견 교환을 통해서 바로 잡도록 하겠다.

참고문헌

- 권세혁. 1999. "웹상에서 통계정보 시스템의 구축." 《한국 데이터 정보과 학회 학술 발표회 논문집》, 7-16.
- 이계오. 1998. "지역 통계 발전 방안." 《지역 통계의 활성화 방안 세미나 논문집》 39-67
- 이계오. 1999. "소지역 추정법을 이용한 충북 시군구의 실업자 추정." 《소 지역 통계 워크샵논문집》 37-56
- Couper, Mick. P. 1998. 《Computer Assisted Survey Information Collection》 John Wiley & Sons, New York
- Dillman, Don. A. 1999. "Principles for Constructing Web Surveys."
- Farmer, Tregg. 1999. "Using the Internet for Primary Research Data Collection" Market Research Library.
- Schaefer, David R. 1998. "Development of A Standard E-Mail Methodology" 《Public Opinion Quarterly》 Vol.62, 378-397
- Swoboda, Walter J. 1997. "Internet Surveys by Direct Mailing" 《Social Science Computer Review》 Vol.15, 242-255