

컴퓨터 이용 통계조사의 현황과 전망

Current Situation of the Computer-Assisted Statistical Surveys and Its Prospects

이 계 오
Lee, Kay-O

지식 정보화 사회에서 필요한 통계는 “기술통계+추론통계+실용통계”의 특성을 갖추어야 하며 적시성, 정확성과 경제성을 갖는 통계를 생산하기 위해서 자료 수집과 분석 과정에서 컴퓨터를 이용해 왔으나, 앞으로 인터넷이 보편화된 사이버 공간 사회에서 활용될 수 있는 실용통계 생산에 역점을 둔 통계 발전 방안을 살펴보았다. 먼저 컴퓨터 보급과 이용 현황을 살펴보고 컴퓨터 이용 조사 기법의 발달 과정을 요약하였으며, 적용 사례 중심으로 인터넷 조사의 발전 가능성에 대해서 언급한 후에, 모든 분야에서 어떤 종류의 통계도 효과적으로 생산할 수 있는 컴퓨터 이용 통계 조사 종합 시스템을 제안하였다.

주요용어 : 실용통계, CASIC, 통계조사 종합시스템

In this paper, the following facts have been reviewed:

First, statistics should have the characteristics of descriptive, inferential, and applicative statistics in "Information-Knowledge Society." Second, the computer has been employed in the process of collection and data analysis for producing the statistics with timeliness and efficiency. Third, we have to develop a statistical development plan which lays emphasis on the production of applicative statistics in the coming "Internet Cyber Society."

In addition, we investigated overall status of computer utilization, and the process of the development of computer-assisted statistical survey methods. The prospects of internet surveys were also investigated here. On the basis of the above review and investigation, we suggested "an integrated system of computer-assisted statistical surveys" that can be applied to all kinds of surveys and statistics.

1. 서론

새로운 천년의 시작인 21세기에서 국력의 원천은 지식과 정보를 기반으로 할 것이며 지식의 특성과 유용성이 변하고 있다는 말을 우리는 많이 들어 왔다. 특히 지식과 정보의 중요성을 강조하는 지식 정보화 시대의 도래를 이미 조금씩 체험하고 있을 것

이다. 지식과 정보는 독립적인 개념이 아니고 상호 연계되어 작용하고 영향을 주고 있으며, 앞으로 정보 통신 기술과 사회 생활 환경의 변화에 따라 상호 작용이 더 활성화될 것이다. 생활 환경과 의식의 변화에 따라서 지식의 개념과 유용성도 달라지고 있으며 특히 지식의 활용성과 내용에 따라서 지식을 事物知(Knowledge about Objects), 事實知(Knowledge about Propositions)와 方法知(Knowledge about Know-how)로 분류하고 있으며 정보화 사회로 발전하면 할수록 方法知의 영향이 증대될 것이다. 위에서 事物知는 실체의 존재 가치를 인식하는 개념적인 지식이고 사실지는 현상과 진리를 탐지하는 지식이며 방법지는 인간의 욕구를 충족시킬 수 있는 방법을 아는 지식으로 창조적 가치 창조가 중요시되는 정보화 사회에서는 방법지의 영향이 확대될 것이다.

지식 정보 사회에서 지식 정보의 기초 요소인 통계의 특성과 갖추어야 할 조건들을 살펴보자. 통계는 지식 정보의 인프라적인 역할을 하고 있으나 중요성에 대한 인식은 미흡할 뿐 만 아니라 잘못 이용되면서 부정적인 평가까지도 받고 있다.

특히 21세기에는 모든 지식이나 기술 환경의 변화가 빠르게 진전되므로 시의 적절한 통계의 생산 공급과 활용은 필수 불가결한 요소가 될 것이다. 앞으로 통계는 지금까지 통용되어 왔던 기술 통계(Statistics about Objects)와 추론 통계(Statistics about Propositions)와는 다른 특성을 갖는 實用 統計(Statistics about Know-use)의 내용을 갖춘 것이어야 할 것이다. 물론 지식 정보화 시대에 요구되는 통계의 기능을 다하기 위해서는 “기술 통계 + 추론 통계 + 실용 통계”의 특성을 함께 갖추어야 하겠지만 21세기에서 필요한 가치 창조와 경쟁력 제고 차원에서 본다면 실용 통계의 개발 연구에 더 많은 비중이 주어져야 할 것이다. 실용 통계가 갖추어야 할 기본적인 속성으로는 신속성, 정확성과 경제성을 생각할 수 있으며 신속성이란 사회 의식과 환경의 빠른 변화에 대처하고 정책 입안이나 기업의 의사 결정의 기초 정보로 활용될 수 있는 최신의 통계 정보로써 특성을 말하며, 정확성이란 모집단의 모수를 얼마나 잘 대변할 수 있는지의 측도를 말한다. 기본적인 3가지 속성은 상호 반작용적인 특성을 갖고 있으므로 최선의 방안을 찾는 것은 쉽지 않을 것이므로 최적의 방안을 모색하는 것이 바람직할 것이다.

기존의 통계 조사 기법인 우편 조사, 전화 조사 또는 면대면 면접 조사법에서 3가지 기본 속성을 개선하기 위한 노력으로 컴퓨터를 이용한 통계 조사 기법이 시작되었다. 컴퓨터 기술의 발달과 통계 조사 이론의 발전으로 실용 통계의 생산과 활용에서 많은 진전이 있었으므로 현재 시점에서 정보 통신 기술의 발달을 감안하여 미래를 전망하기 위해서 먼저 컴퓨터 이용 현황을 알아보고 컴퓨터 이용 조사 기법의 발달 과정을 요약하며, 인터넷 조사의 이용 사례를 살펴본 후에 컴퓨터 이용 통계 조사 체계의 발전 방안을 제안하고자 한다.

2. 컴퓨터 이용 현황

실용 통계의 기본 속성인 신속성, 정확성과 경제성을 제고하기 위한 통계 생산 방법으로 컴퓨터를 이용하게 됨에 따라 컴퓨터의 이용 현황을 살펴봄으로써 모집단을 대표하는 표본을 추출하고 정확하게 조사할 수 있는 여건을 갖추었는지를 판단할 수 있을 것이다.

먼저 조사 단위인 가정의 컴퓨터 보급과 이용 실태를 한국정보문화센터의 자료를 중심으로 살펴보자. 작년 후반에 시작된 국민 PC보급 정책으로 금년 5월 기준으로 66%의 가정에서 컴퓨터를 보유하고 있으며 컴퓨터 기종도 펜티엄급 이상이 82%로 가정의 컴퓨터 보급률은 선진국 수준이다. 연도별 컴퓨터 보급률은 <표 1>에 주어졌으며, 가정 보유 컴퓨터 기종별 현황은 <표 2>에 주어졌다.

<표 1> 가정의 컴퓨터 보급률

연도	1997	1998	1999	2000	미국	싱가폴
보급률	43.2	44.5	51.8	66.0	90.0	95.0

<표 2> 컴퓨터 보유 기종(%)

기종	386급 이하	486급	펜티엄 이상	기타
1999년	6.8	26.2	64.4	2.7
2000년	1.8	14.2	81.8	2.3

가정에서 컴퓨터 보유도 중요하지만 통계 조사에서는 컴퓨터 이용률과 이용자들의 연령대별 분포도 중요한 사항이므로 이용률을 알아보면 국민의 반 이상인 50.4% 정도가 이용하고 있을 뿐 만 아니라 이용 장소도 직장이나 학교뿐만 아니라 가정에서 이용도 많은 것으로 나타나고 있다. 사회 조사에서 참고 사항으로 필요한 인구 사회학적인 변인별 이용률을 보면 남자가 58.7%, 여자는 41.9%이고 연령대별로 10대가 91.7%로 가장 높고 다른 연령대별 이용률은 <표 3>에 주어졌다.

<표 3> 연령대별 이용률 및 이용시간

연령대별	10대	20대	30대	40대	50대 이상
이용률(%)	91.7	71.7	46.1	31.6	10.2
이용시간(분)	147.0	234.9	192.7	183.4	144.0

컴퓨터 이용 시간은 하루 평균 191.7분으로 3시간 넘게 컴퓨터를 이용하고 있는 것으로 조사되었으며 이용장소별로는 가정에서 108.5분, 직장이나 학교에서는 185.3분으로, 업무 처리나 공부 관계로 많이 이용하고 있다. 연령대별 사용 시간은 20대가 234.9분으로 4시간 정도 이용하고 10대는 이용률은 높지만 이용 시간이 길지 않은 것으로 조사되었다. 직업별 이용 시간은 <표 4>에 주어졌으며 화이트칼라가 287.1분으로 4.8시간 이용하고 대학생이 232.8분으로 4시간 정도로 두 번째로 길게 이용하며 아직까지는 주부의 이용 시간이 88.2분으로 가장 짧은 시간으로 나타났다.

<표 4> 직업별 컴퓨터 이용시간(분/일)

직업	농림업	자영업	블루칼라	화이트칼라	주부	중고생	대학생
이용시간	88.2	194.8	166.3	287.1	87.4	138.5	232.8

컴퓨터 이용자들의 이용 시간은 저녁 8시부터 11시까지가 가장 많고, 다음으로는 오전 10시부터 12시까지 많이 이용하며 오후 3시부터 5시까지도 많이 이용하는 현상을 보여 이용 패턴이 오전, 오후와 저녁 시간대에 한 번씩 많이 이용하는 시간대가 있는 형태로 주어졌다. 이는 컴퓨터의 이용이 실제 생활과 밀접한 관계가 있으며 컴퓨터가 생활의 도구이자 기기로 자리잡아가고 있음을 보여주고 있다.

<표 5> 연령대별 컴퓨터 이용 목적 (%)

구 분		공부나 연구 활용	회사/직장 업무처리	자격취득 취업준비	여가시간 활용	정보관리 (개인/가정)	기타
성 별	남자	16.8	31.9	4.6	41.1	5.0	0.5
	여자	20.5	23.7	6.9	37.1	11.0	0.8
연 령 대 별	10대	27.0	4.5	11.0	54.7	2.9	0.0
	20대	23.0	29.7	5.4	35.1	6.6	0.2
	30대	7.7	45.7	2.6	30.3	12.3	1.5
	40대	8.6	44.3	0.6	33.3	11.5	1.7
	50대이상	5.8	42.3	0.0	42.3	7.7	1.9
전 체		18.3	28.6	5.6	39.5	7.5	0.6

생활 속에 필수품으로 자리를 잡아가고 있는 컴퓨터 이용 목적은 성별과 직업에 따라서 상이하게 주어졌으며 전체적으로 39.5%가 여가 시간의 활용을 위해 컴퓨터를 이용하며 다음으로는 직업상 업무처리 용도로 28.6%가 사용하고 있고, 공부 또는 연구를 위해서는 18.3%가 컴퓨터를 이용하고 있다.

성별로 이용 목적을 보면 남자들은 회사/직장 업무처리 용도로, 여자들은 공부나 연구를 위해서 상대적으로 컴퓨터를 많이 이용하고 있다. 연령대별로 10대와 20대는 여가 시간 활용을 위한 용도로 많이 사용하고, 30대와 40대는 회사/직장의 업무 처리를 위해서 주로 이용하며 50대 이상은 여가 시간 활용을 목적으로 컴퓨터를 이용하는 것으로 나타났다.

정보화 시대에서 정보 교환의 기간망인 인터넷에 관한 언급없이 21세기의 통계 조사의 전망을 말할 수 없을 것이다. 정보통신기술의 발달과 컴퓨터의 보편화로 지구촌이 정보의 단일 생활권으로 변하게 되면 통계 조사 환경과 기법들이 달라져야 할 것이다. 먼저 우리의 인터넷 이용 현황을 알아보자.

먼저 인터넷 이용률을 보면 13세에서 49세까지를 대상으로 한 인터넷 이용률이 48.6%로 세계적으로 상위권에 속하고 있으며 이용자 수도 1420만 명이 된다. 연령대별 이용률과 이용자 수는 <표 6>에 주어졌다.

<표 6> 인터넷 이용률과 이용자 수

구분	10대	20대	30대	40대	전체
이용률(%)	66.9	64.5	39.8	23.8	48.6
이용자수(천명)	3,620	5,536	3,534	1,519	14,208
구성비(%)	25.5	38.9	24.9	10.7	100.0
이용시간(분)	74.4	115.2	110.1	94.3	99.5

이용자수는 20대가 554만 명으로 제일 많고 이용률은 10대가 66.9%로 제일 높으며 40대의 이용률은 23.8%일 뿐 만 아니라 증가 추세가 정지된 현상을 보이고 있다. 직업별 이용률에 대한 자료는 <표 7>에 주어졌으며 직업간의 격차가, 농/임/어업 7.3%에서 대학생 94.1%로 ,86.8% 정도 상당히 크게 나타났다.

<표 7> 직업별 이용률 및 이용시간

구분	농/임/어업	자영업	블루칼라	화이트칼라	주부	중고생	대학생	기타
이용률(%)	7.3	18.7	27.8	64.0	11.0	72.7	94.1	43.4
이용시간(분)	49.1	101.2	87.6	141.0	55.3	68.6	119.5	81.5

특히 중고생과 대학생의 이용률이 높게 나타났으며 화이트 칼라의 이용률도 64.0%로 높지만, 농/임/어업은 7.3% 아주 낮아서 도시와 농어촌 지역 간의 인터넷 이용률의

격차가 큰 것으로 판단된다. 앞으로 정책적으로 농어촌 지역의 정보화 사업이 추진된다면 도시와 농어촌간의 인터넷 환경의 격차는 좁혀질 것으로 기대된다.

하루 이용 시간을 살펴보면 국민의 평균 이용 시간은 99.5분으로 1시간 30분 정도이며, 연령대별로는 20대와 30대가 115.2분과 110.1분으로 하루 평균 2시간 정도 인터넷을 이용하고 있으며, 직업별로는 화이트 칼라가 141분으로 2시간 20분 정도 이용하고 다음으로는 대학생이 119.5분으로 2시간 정도 인터넷을 이용하고 있는 것으로 나타났다. 직업별 이용 시간의 격차는 농/임/어업이 49.1분으로 가장 짧고 화이트 칼라가 141분으로 장시간 이용하여 두 직업간의 차이가 92분이나 된다.

인터넷 이용 시간대는 오전보다는 오후, 오후보다는 저녁 시간에 많이 이용하고 있으며, 밤 9시에서 12시 사이가 가장 많이 이용하는 시간대이므로 인터넷 접속률이 낮은 지역에서는 참고할 내용이다. 인터넷 이용 목적에서는 남자와 여자가 모두 여가 시간 활용과 생활/취미 정보 획득의 비중이 크며, 연령별로는 10대와 50대의 경우 여가 시간 활용과 의견/메시지 교환이 상대적으로 많은 반면, 20대는 업무/사업 투자 관련 이용이 상대적으로 많았고, 30대와 40대는 생활/취미 정보 획득과 업무상 의견/문서 교환이 상대적으로 많았다. 인터넷 공간이 생활 정보 교환과 업무 활동 영역으로 확대되어 감에 따라 인터넷을 이용한 통계 조사는 새로운 연구 개발 분야가 될 것이므로 다음에 자세하게 다루겠다.

3. 컴퓨터 이용 조사 방법

컴퓨터의 대중화와 가정의 보급률이 높아짐에 따라 컴퓨터를 기반으로 한 조사 방법이 이제는 선택적인 방안이 아니고 중심적인 방법이 되어가고 있다. 컴퓨터 이용 조사 방법의 총괄적인 명칭은 CASIC(Computer Assisted Survey Information Collection)으로 불리질 수 있으며, 사회 의식의 변화와 정보 통신과 컴퓨터 기술의 발달로 CASIC을 생각하게 되었으나 통계의 기본 속성인 신속성(적시성), 정확성과 경제성을 충족하기 위해서는 기존의 통계 조사 방법의 기본 원칙인 모집단 정의, 추출틀 작성, 표본 추출법, 설문지 작성과 자료 입력 및 분석 공표 등도 다시 검토되어 주어진 여건에서 최적의 방법이 강구되어야 할 것이다. CASIC의 시초는 CATI(Computer Assisted Telephone Interviewing)로서 AT&T의 R.M. Gryb에 의해서 제안되었으며, 1971년 처음으로 조사 현장에서 사용되었으나 1978년 Gerald Glasser에 의해서 "CATI"라고 불리워졌다. CATI의 도입 초기에는 P&P(Paper and Pencil) 전화 조사에 비해서 비용 절감, 시간 단축과 자료의 질적 개선 등을 주장했으나 시장 조사의 실험적인 조사에서 검증되지 않아서 논쟁이 있어 왔으며, CATI에 맞는 조사 설계와 추정 기법 등이 개발되고 컴퓨터의 속도와 용량이 발달하면서 CATI의 사용에 대한 논란 등은 없어지고, 1980년 후반 이후부터 시장 조사와 사회 의식 조사 또는 선거 여론 조사 등에서는 CATI에 의한 통계 조사가 큰 비중을 차지하고 있다. 1990년 초반까지

도 CATI에 의한 통계 조사가 가장 보편화된 조사 기법으로 자리잡았고 우리나라에서는 현재도 CATI의 유형에 속하는 전화 조사 기법이 통용되고 있다.

CATI는 중앙 통제 형식의 전화 조사 방법으로 이용되었으나 정부 등에서 실시하는 대규모의 광범위한 지역 조사에서는 각 지역별로 분할 관리 체계가 적용되었으며, 면대면 P&P 면접 조사에서는 CATI보다는 CAPI(Computer Assisted Personal Interviewing)이 1970년 후반부터 이용되었다. 특히 면접조사에서 면접원에 의한 비표본 오차의 발생을 줄이고 민감한 사안에 대해서는 응답자 본인이 질문을 읽고 키보드에서 응답하는 CAPI가 더 효과적이므로 널리 이용되었으며, 특히 반도체 기술의 발달로 휴대용 컴퓨터가 경량화되고 비용이 저렴해지며 S/W가 개발되면서 1980년대부터 미국뿐 만 아니라 서부 유럽에서는 CAPI에 의한 사회 의식 면접 조사가 보편화되기 시작하였다. 특히 1987년 독일에서 노동력 조사에 CAPI를 대규모 국가적인 조사에 적용하면서부터 미국, 영국, 캐나다, 프랑스 등에서도 노동력 조사에 CAPI를 이용하여 자료 입력, 검색 및 자료 처리에서 시간 단축을 가져왔다.

소비자 물가 지수 조사를 CAPI를 이용하기 위해서 마이크로 컴퓨터에 바코드 스캐너를 설치하여 표본 가구가 구입한 물건의 코드와 바코드화된 질문의 응답을 읽도록 하여 응답자의 부담을 감소함으로써 조사 자료의 질을 향상시키기도 하였다.

우리 나라에서 매일 실시하는 경제 활동 인구 조사에서 노트북을 이용한 CAPI 조사 기법을 1998년부터 부분적으로 적용하여 왔다. 민감한 사안을 조사하거나 응답자의 대답 내용에 따라 질문 순서를 조정하는 조사에서는 설문 내용과 조사 과정을 담은 디스켓을 표본 가구 또는 조사 대상자에게 전달한 후 응답자는 디스켓을 자신의 컴퓨터에 설치하여 조사 과정을 완료한 후에 조사 센터로 디스켓으로 보내거나 e-mail 등으로 조사 본부에 보내는 형식을 CSAQ(Computerized Self-Administered Questionnaire)라고 한다. CSAQ의 형식은 조사 목적, 조사 내용에 따라서 구체적인 기법이 다르기 때문에 일괄적으로 언급하기 어렵지만 응답자들의 PC 운영 체제와 성능 및 S/W에 따라서 다양한 방법을 적용할 수도 있으나 제한을 받을 수도 있다. CSAQ를 일반적인 모집단에 적용은 어렵지만 범위 편향을 줄이기 위해서 통상적인 방법으로 표본을 선정하고 PC가 없는 응답자에게는 PC를 대여해주는 방안이 이용되기도 한다. 컴퓨터를 이용한 자료 수집 방법으로는 CATI, CAPI와 CSAQ 이외에도 민감한 사안 등을 조사할 때 적용가능한 기법인 CASI (Computer Assisted Self-Interview)이 있으며 CASI는 랩탑에 조사 전용 S/W를 내장하여 응답자가 모니터를 통해서 질문을 읽고 키보드로 응답하는 형식이며 때로는 P&P 면대면 면접 조사를 하다가 민감한 내용을 물을 때는 응답자의 정확한 대답을 얻기 위해서 CASI를 이용할 수도 있다. 시각 장애자 또는 글을 제대로 읽지 못하는 응답자에게는 Audio CASI 기법을 이용할 수 있는데 ACASI는 응답자가 헤드셋을 쓰고 모니터로 질문을 읽음과 동시에 컴퓨터 음성으로 질문을 듣고서 키보드로 대답하는 형식이다.

TDE(Touch-tone Data Entry)는 조사 센터로 연결된 무료 전화 번호를 이용하여 응답자들이 전화를 걸어서 컴퓨터 시스템에서 읽어 주는 질문에 대한 대답을 전화기

keypad로 입력하는 형식이며, 1987년 미국 노동성 통계국에서 경상 취업 조사에 처음 도입한 이후로 기업체 조사 또는 여론 조사 등에 이용되고 있으나 최근에는 TV 방송 등에서 ARS 방식으로 즉흥적인 여론 조사를 하고 있으나 표본 대표성에 문제가 있을 수 있다.

컴퓨터 입력 시스템의 발달로 TDE의 대체 방안으로 VRE(Voice Recognition Entry)가 제한적으로 이용되고 있다. VRE는 TDE와 유사하지만 응답시 keypad를 사용하지 않고 음성으로 대답하는 형식이며, 아직까지는 복잡한 음성 응답의 인식에 문제가 있으므로 간단한 대답인 “예”, “아니오” 또는 숫자와 기호화된 대답만 받아들일 수 있으며 반드시 대답이 끝난 후에는 응답 내용이 컴퓨터에 제대로 입력되었는지 확인하는 절차를 거쳐야 정확한 조사가 이루어질 수 있다. 앞으로 컴퓨터 입력 기술이 발달한다면 민감한 내용이나 정성적인 조사까지도 면접원없이 실시할 수 있을 것으로 기대된다. 지금까지는 조사 과정에서 컴퓨터를 이용하는 방법들을 알아보았으나 다음에는 자료 편집 과정에서 컴퓨터를 이용하는 기법을 살펴보자.

CAC(Computer Assisted Coding)은 과거의 수기식 코딩에서 발생할 수 있는 비표본 오차를 감소시키기 위해서 컴퓨터 내부에 거주지, 직업, 취업 산업 분야, 생산품 등을 사전에 입력하여 코드화한 후에 응답을 코드로 하도록 하는 기법이다. 다음에는 기업체 또는 사업체들의 업무가 전산화되어감에 따라 통상적인 업무처리 내용들로부터 경제 관련 자료 또는 가격 등의 정보를 자동 정보 전환 프로그램을 이용하여 얻어내는 방법이 연구되고 있으며 이를 EDI(Electronic Data Interchange)라고 한다. 앞으로 EDI를 활용할 수 있는 환경이 갖추어진다면 통계 조사를 직접할 필요성이 감소되고 행정 업무 자료와 기업들의 경제 활동 표준화를 통해서 필요한 통계를 작성할 수도 있을 것이다.

이상에서 설명한 내용을 총괄적으로 포함하여 CASIC으로 나타내기도 하지만 아직까지 공통적으로 사용하는 용어로 인정되지 않았지만 본 고에서는 그대로 사용하였다. 모든 조사에서 CASIC이 신속성(적시성), 정확성과 경제성을 갖는 것은 아니고 조사 내용이나 조사 규모에 따라서 CASIC의 영향이 다를 수 있다. 일반적으로 CASIC의 적용은 자료 수집과 분석에서 시간을 단축할 수 있고, 비용 절감의 효과를 얻을 수 있으며, 무응답 항목과 편집 오류의 감소로 자료의 질을 높일 수 있다는데는 공감할 것이다.

앞으로 컴퓨터를 이용한 조사 기법이 개발될지라도 전통적인 조사 기법인 P&P 면접이나 설문지를 이용한 우편 조사들도 함께 발전하여 상호간의 문제점을 보완하는 이들의 혼합 조사 방법이 활용될 것으로 기대된다. 혼합 조사 방법을 운용할 수 있는 시스템을 발전 방안에서 자세하게 언급할 것이다.

4. 인터넷 조사법

우리나라의 인터넷 이용자가 1500만 명에 달하고, 인터넷에 접속된 기관, 단체와 개인의 성격이나 규모에 제한이 없을 뿐 만 아니라 지역적인 경계가 없으므로 정보의 유통 속도가 빠르고 또한 획득에서도 경제적 부담이 적기 때문에 사회 의식 또는 여론 조사에서는 고려되어야 할 조사 방법이다. 먼저 인터넷 조사의 개념을 알아보고, 다음으로 인터넷 조사의 활용 실태와 전망에 대해서 살펴보자.

4.1. 개념

인터넷 조사는 1980년대 초에 특정 집단을 대상으로 인트라넷을 이용한 전자 우편 조사를 통해서 활용 가능성이 입증된 후에 조사 기간 단축과 자료 입력 및 처리 과정에서 정확성을 제고할 수 있는 장점 때문 꾸준한 발전을 해왔으나, 1990년 초반까지는 크게 각광을 받지 못하다가 WWW 사용이 확대되면서 인터넷 이용이 간편해지고 이용자가 급증하게 되면서 인터넷 조사가 활성화되었다.

인터넷 조사의 정의는 인트라넷을 포함한 인터넷의 망에서 이루어지는 조사를 총괄적으로 말하며, 조사 대상자의 선정 방법과 이용 S/W의 수준에 따라서 인터넷 조사를 분류하지만 응답자들의 참여 형식에 따라서 방문조사(Visitor Survey)와 메일 조사(Mail Survey)로 분류할 수 있을 것이다. 방문자 조사는 조사 기관이 홈페이지 등의 포털 사이트에 설문 내용을 게시하고 자발적인 참여자들이 설문에 응답하도록 하는 형식이므로 응답자들 중에서 적극 찬성하는 사람과 적극 반대하는 사람들의 구성비가 높을 수 있으며 중간 계층 또는 관심이 적은 사람들의 의견 반영 폭이 낮을 수 있는 문제가 있다. 메일 조사는 조사를 주관하는 기관이나 개인이 설문지를 작성하여 인터넷을 통해서 조사 대상자들에게 설문지를 전송하고 응답자들은 자신의 컴퓨터에서 설문지에 대답하고 그 결과를 조사 주관자에게 인터넷을 통해서 전달하는 형식이다. 물론 방문자 조사와 메일 조사의 중간 형태 또는 인터넷 조사와 기존의 조사 기법을 혼합한 형태를 생각할 수 있으나 본고에서는 더 이상 언급을 하지않겠다.

앞에서 생각해본 인터넷 조사의 특징은 조사비용이 저렴하고 조사기간이 단축되며 자료입력과 분석에서 비표본 오차를 감소시킬 수 있는 장점을 들 수 있다. 조사내용에서도 인터넷 조사를 위한 시스템 설계와 프로그램개발 비용 등을 생각한다면 통상적인 우편조사나 전화조사보다 비용이 절감되지 않을 수 있고, 유럽 등 과 같이 인터넷 접속료와 사용료가 비싼 경우에는 비용경감의 효과가 적거나 없을 수 있다. 많은 실험적인 조사를 통해서 검증한 내용은 인터넷 조사비용은 전화조사의 반정도 소요되고 우편조사의 1/5정도로 저렴하다는 것이다. 조사소요기간도 인터넷 조사는 짧게는 몇 시간에서 1주일정도로 소요될 수 있지만 전화 또는 우편엽서로 독촉하거나 정해진 기간 내에 응답이 없을 경우에 재차 전송하는 방법을 함께 사용한다면 응답률도 높이고 조사기간도 단축시킬 수 있을 것이다. 자료입력과 분석과정에서 정확성 제고는 우편조사 또는 면접조사의 수기식 작업에 비교하면 쉽게 이해되지만 컴퓨터 이용한 전화조사 또는 면접조사(CAPI)와 비교한다면 별로 차이가 없을 것이다.

인터넷 조사의 문제점은 모집단을 대표하는 조사대상을 선정하는데 어려움이 있다는 것과 응답자의 식별이 어렵고 중복응답에 대한 검색이 쉽지 않다는 점이다. 특히 방문자 조사에서 의도적으로 중복응답을 시도할 경우에는 Cookie정보를 이용하여 중복응답을 검색하는 것도 효과가 없을 것이므로 정확한 자료수집을 위해서 해결해야 할 과제이다.

S/W 수준에 따른 조사 방법의 분류는 Flat 파일, Interactive 형식과 On-line Chat 방식으로 나눌 수 있으며 이들의 장단점은 정보통신기술과 인터넷 S/W의 발달로 달라질 수 있을 것이지만 현재 수준에서 알아보면 다음과 같다.

Flat 파일 형식은 Web 페이지에 설문 내용을 게시하거나 설문지를 인터넷을 통해서 e-mail로 전송하고 수신하는 방법으로 현재 네티즌들에게 익숙한 형태이며 S/W 개발비용도 저렴하고 간단하여 많이 이용되고 있지만 질문 형식과 문항 수 등에서 제한을 받는 단점도 있다. Interactive 형식은 전용 서버를 사용하고 Interactive 프로그램을 통해서 연결된 대상자들간의 조사에 이용되는 방식으로 CATI 또는 CAPI를 연계해서 조사할 수 있고 질문 순서의 조정과 앞에서 응답한 내용을 참조하여 뒤에 오는 질문에 대답할 때 이용할 수도 있다. 전송 속도와 신뢰성 문제와 보안성의 약점을 해결한다면 앞으로 인터넷 조사의 주된 방법으로 발전할 것으로 생각된다.

On-line chat 형식은 네트워크로 연결된 컴퓨터를 이용하여 개인간의 메시지를 실시간으로 상호교환하면서 대화를 통하여 조사를 수행하는 방법이며, 현재는 네트워크 시스템의 제한과 인터넷 용량과 속도상의 한계로 제한적으로 사용되고 있으나 정성적인 특성 연구 등에 활용되고 있으며, 앞으로 집단 면접 조사의 대체 방법으로 유망한 형식이다. 특히 동화상 등의 다양한 시청각 보조 장비를 이용한다면 복잡한 내용을 간편하게 조사할 수 있어서 인터넷 조사의 차세대 기법이 될 전망이 높다.

인터넷 조사를 계획할 단계에서 다음 3가지 사항에 유의해서 결정해야 한다. 먼저 조사 대상 모집단과 조사 가능 모집단간의 차이가 없는가? 만일에 차이가 있다면 모수 추정에서 차이를 보정할 수 있는 추가 정보 또는 추정법이 있는지를 검토해야 한다. 다음에는 Web에서 게시할 질문과 자료 내용에 비밀 사항이나 개인적으로 민감한 사항은 없는지를 확인하고 만일에 있다면 비밀 사항을 보호할 수 있는 보안성을 갖춘 중앙 통제 방식의 조사 방법을 적용하는 방안을 고려할 수 있을 것이다. 마지막으로 조사 소요 기간과 질문 분량은 적절한지 검토해보고 질문은 간단하게 조사 소요 시간을 짧게 할 수 있다면 이는 응답률을 높이고 조사 자료의 질을 높이는데 도움이 될 것이다.

4.2. 인터넷 조사 이용 실태

통계의 특성인 적시성과 정확성을 충족시킬 수 있는 조사 기법의 하나가 인터넷 조사가 될 가능성을 보여준 실제 조사 사례를 알아보자. 먼저 영국에서 1997년 5월 총선 여론 조사에서 UKOPINION 회사가 실시한 인터넷 조사를 보면 조사 참여자를 신

문 광고, 포털 사이트 광고와 회사 전용 패널 구성원으로부터 인센티브를 내걸고 모집하였다. 이 당시 영국의 인터넷 이용률은 10%를 약간 상회하는 수준이었으며 조사 참여자는 600명이었다. 투표 전에 인터넷 여론 조사를 3월 하순과 4월 하순에 2회 동일 참여자를 대상으로 실시하였다. 3월의 조사에서는 590명의 응답자를 확보하였으며, 4월 조사에서는 450명에게 설문을 전송했으나 201명 만의 유효 응답을 수집하여 분석하였다.

인터넷 조사 결과와 기존 조사 결과의 비교가 <표 8>에 주어졌다.

<표 8> 인터넷과 기존 조사 방법의 비교(%)

구분	인터넷 조사		기존 여론 조사		투표 결과
	3월	4월	3월	4월	
보수당	26.0	32.0	29.5	31.5	31.4
노동당	52.6	46.5	52.0	47.5	44.4
자유민주	9.0	13.5	12.5	14.0	17.2
개혁당	6.0	3.0	.	.	2.4
기타	6.5	5.0	6.0	7.0	4.6

인터넷 조사 참여자의 성별 분포는 남자 87%, 여자 13%로 남자에 편중되었으며, 연령대 분포는 17~24세: 17%, 25~34세: 30%, 35~44세: 26%, 45~54세: 18%, 55세 이상: 9%로 44세 이하가 83%로 젊은 층에 편중된 현상을 보이고 있다. <표 8>에서 인터넷 조사와 기존 여론 조사간의 차이가 거의 없음을 볼 수 있으며 노동당은 인터넷 조사와 기존 여론 조사에서 상향 추정됨을 알 수 있고 보수당은 인터넷 조사와 기존 여론 조사의 추정값과 투표 결과가 거의 유사함을 볼 수 있다. 보수주의자와 진보주의자의 사전 여론 조사에 응답하는 성향에 차이가 있는 것으로 볼 수 있으며 일반적으로 보수 성향의 지지 기반은 40대 이후이므로 인터넷 조사에서 응답자가 극히 소수일 수 있고 또한 보수 성향 의식을 갖는 사람들은 여론 조사 등에 응답을 회피하거나 정확하게 응답하지 않을 가능성이 큰 점에 유의해야 한다.

우리나라에서도 2000년 4·13 총선 여론 조사에서 한국 갤럽에서 실험 조사의 일환으로 Daum의 홈페이지에 조사 참여자를 모집하는 광고를 게시한 후에 인터넷 선거 여론 조사에 참여한 사람을 대상으로 해당 선거구의 지지 후보를 묻는 형식으로 조사를 실시하였다. 전국적으로 20,575명의 유효 응답자를 확보하였으나 선거구별로 분석했을 때 어느 정도 신뢰성을 갖춘 추정값을 얻기 위해서 200명 이상의 응답자가 있는 선거구 5군데만을 분석하였으나 서울 지역의 4개 선거구의 인터넷 조사 자료만을 살펴보자. 추정 방법은 성별-연령대별로 불균형된 분포 형태를 갖기 때문에 유권자 구성비를 이용하여 가중값을 부여한 추정값을 계산하는 형식을 이용하였으며, 전화 조

사 결과와 투표 결과를 비교한 내용이 <표 9>에 주어졌으며 전화 조사 결과와 인터넷 조사 결과를 결합하여 추정값의 정도를 높이려는 시도를 하였으며 이 추정값도 함께 <표 9>에 포함되었다.

<표 9>에 나타난 내용을 요약하면 첫째로는 전화 조사의 표본 수는 인터넷 조사보다 2배 정도 많은데도 추정값의 정도에서 큰 차이를 보이지 않고 있음을 볼 수 있다. 두 번째로는 인터넷 조사에서 민주당 후보의 지지율은 과다 추정되고 있으며 특히 후보자의 연령 차이가 클수록 인터넷 조사에서 과다 추정되는 정도가 크다. 예를 들면 성동구에서 민주당 후보의 인터넷 조사 추정값은 개표 결과보다는 18%정도, 전화 조사보다는 25.6%를 과다추정하였다. 세 번째로는 결합 추정값을 인터넷 조사와 전화 조사의 추정량의 분산을 가중값으로 계산하였으며, 개표 결과와의 차이를 가장 작도록 조정된 것으로 나타났다. 인터넷 조사의 개선 방안으로 연구될 수 있는 방법이 혼합 조사 방법(Mixed Mode)을 고려할 수 있음을 보여준 사례이다.

<표 8> 인터넷 조사와 전화 조사의 비교

구분		한나라당	민주당	자민련	민국당
1강남갑 (262) ^① 466	개표결과	56.5	36.8	3.1	0.8
	②결합추정	58.4	34.2	1.1	0.6
	전화조사	53.8	31.2	6.2	2.1
	인터넷조사	55.6	38.7	0.4	0.4
관악갑 (212) 415	개표결과	29.9	43.9	17.1	0.6
	결합추정	23.3	49.7	16.3	0.5
	전화조사	26.3	36.9	22.5	3.1
	인터넷조사	22.7	51.6	14.6	0.3
성동 (231) 551	개표결과	38.9	48.4	2.3	0.4
	결합추정	29.0	58.2	1.5	2.5
	전화조사	37.6	40.4	6.5	1.4
	인터넷조사	21.4	66.0	1.2	0.5
송파갑 (232) 441	개표결과	54.8	37.2	3.7	0.9
	결합추정	55.3	36.5	3.3	0.7
	전화조사	50.1	35.9	7.5	1.9
	인터넷조사	49.4	44.9	0.2	0.2

① 괄호 속의 숫자는 인터넷 조사 응답자 수이고 그 아래 숫자는 전화 조사의 응답자 수임.

② 결합 추정값은 전화 조사와 인터넷 조사의 추정값들을 결합하여 계산한 추정값임

민간 조사 회사들의 인터넷 이용 조사 사례 이외에도 통계청에서 종합 환경 지수 관련한 조사를 환경 분야 공무원과 교수 및 환경 전문 기자를 대상으로 200명을 조사 대상으로 추출한 후에 전화로 e-mail 주소를 확인하여 e-mail을 통해서 조사 참여를 알렸으며 조사 설문은 통계청 홈페이지에 79개 항목으로 구성된 웹 설문지 형식으로 게시하여 조사하였다. 조사 응답률이 81.5%로써 특정 집단(인터넷 이용률이 높은 계층)을 상대로한 통계 조사로 인터넷 조사가 활용될 수 있음을 보여준 사례이다. 본 조사에서도 공공분야 공무원과 교수들의 응답률은 90% 이상이었으나 환경전문가와 민간환경단체는 60% 이하의 응답률을 보여 계층별로 인터넷 이용 수준의 차이가 있으며 이로 인하여 응답률에서 큰 차이가 있음을 짐작할 수 있다. 물론 인터넷 이용 수준만이 응답률의 차이에 대한 원인이 아니고 직업 의식이나 통계 조사에 대한 인식의 차이가 원인이 될 수도 있기 때문에 좀 더 심층적인 분석이 필요하다. 지금은 인터넷 조사가 초기 단계이므로 기본적인 이론과 기술 연구에 많은 노력과 재원을 투자하여 새로운 인터넷 시대에서는 통계 선진국으로 발전할 수 있는 밑바탕을 마련해야 할 것이다.

4.3. 인터넷 조사의 발전 방안

인터넷 조사에서 선결해야할 과제가 대표 표본의 확보와 중복응답방지 및 응답자의 식별 문제임을 앞에서 언급하였으며, 이들을 해결하고 앞으로 인터넷 이용률이 70% 이상 되었을 때를 대비하여 인터넷 조사를 통해서 적시성과 정확성을 갖춘 통계를 생산할 수 있는 방안을 살펴보자.

먼저 영국의 총선 여론 조사 사례에서 인터넷 조사의 추출틀로서 가장 적합한 것은 인터넷 조사 전용 패널임을 보았다. 조사 목적에 따라 모집단이 결정되므로 모집단의 인구사회학적 구조 분석을 통해서 조사 대상 모집단을 대표할 수 있는 전용 패널을 구축하여 조사를 한다면 신뢰성있는 통계를 경제적으로 생산할 수 있을 것이다. 특히 전용 패널 구축시 인터넷 사용률은 낮고 보수 성향이 높은 노장년층을 대변할 수 있는 패널 구성원의 모집에 유의해야 하며, 모집단과 전용 패널의 구조가 유사할지라도 유효 응답자들의 구조적 특성을 분석한 후에 적절한 추정 방법을 적용해야 할 것이다. 다음으로는 기존의 조사 방법인 우편 조사, 전화 조사와 면접 조사들을 인터넷 조사와 결합한 조사 방법에 대한 연구가 필요하다. 기존의 조사 방법들은 대부분 유용성이 검증되었으므로 인터넷 조사 결과의 유효성을 비교할 수 있는 상이한 조사 방법을 병행하여 적용한다면 추정에서 안정된 추정값을 계산하는데 유익할 것이다. 예를 든다면 선거 여론 조사에서 초기에는 시간적인 여유가 있으므로 대규모 전화 조사를 실시하여 모집단의 구조적 특성도 파악하고 인터넷 조사에 필요한 정보를 확보한 후에는 인터넷 조사를 실행하여 후보자별 지지율의 변화 추세와 지지율을 추정하는 방법을 생각할 수 있다. 이 때 추정법은 전화 조사와 인터넷 조사를 독립적으로 실행할

경우는 두 추정량의 가중 평균 추정값을 구하는 방법을 적용할 수 있고 전화 조사의 조사 대상자들을 부분적으로 교체하여 일부는 반복적으로 인터넷 조사에서 조사 대상으로 하고 또 일부는 새로운 조사 대상으로 추출하여 조사하는 경우에는 회귀 추정법을 적용하여 후보자별 지지율을 추정할 수 있다. 이와 같이 두 가지 이상의 조사 기법을 혼합하여 동일 주제를 조사하는 방법을 혼합 조사법(Mixed Mode) 또는 다양 조사법(Multi-Mode)라고 한다.

앞으로 혼합 조사법에 관한 내용은 이론적으로나 실제 조사 설계 측면에 심층적으로 연구되어야 할 과제이다. <표 9>에 주어진 결합 추정값은 서로 독립적인 추출틀과 조사 기법을 적용한 혼합 조사법의 추정법의 유용성을 보여 주는 좋은 예가 될 것이다.

5. 컴퓨터 이용 통계 조사 종합 시스템

지식 정보화 시대에서 실용 통계로써 활용되기 위해서 갖추어야 할 특성을 적시성(신속성), 정확성과 경제성이라고 앞에서 언급하였다. 3가지 특성을 동시에 만족하는 조사 기법을 찾기가 어려울 것이므로 주어진 연구 주제와 상황에 따라서 어떤 특성에 우선 순위를 두느냐에 따라서 적합한 조사 기법과 자료 분석 방법이 정해질 것이다. 적합한 통계 조사 기법을 결정하고 자료 수집, 입력과 검색의 과정을 선택하며 자료 분석용 DB 구축 및 분석 결과의 생산까지 Interactive하게 해결하는 통계 조사 종합 시스템이 있다면 적시성, 정확성과 경제성을 충족하는 통계의 생산이 가능할 것이다.

통계 조사 종합 시스템의 유사 유형으로는 네덜란드 통계국에서 개발한 BLAISE를 생각할 수 있을 것이다. BLAISE 시스템은 설문지 설계에서부터 자료 수집, 자료 처리와 통계 작성 과정까지를 단일 시스템에서 처리하도록 설계되었으며, 설문지 설계에서 Blaise라는 언어를 이용할 수 있으며 이는 Pascal 언어와 같은 구조형 언어이며 간편하고 데이터 베이스 구축을 염두에 두었다. 자료 수집 방법으로는 CAPI, CATI와 CASI 또는 기존의 조사 기법인 우편 조사, 면접 조사 등 중에서 한 가지를 선택하거나 2가지 이상을 결합한 형식을 조사 주제와 환경에 따라 선택할 수 있고 설문지 발송과 중간 확인을 통한 재발송 절차 등을 포함한 자료 입력 시스템에는 무응답이나 결측값에 대한 대체 방법을 선택할 수 있는 프로그램이 포함되어 있다. 자료의 변환과 테이블 작성 과정을 포함하고 있고 원자료를 메타 자료와 연계하여 DB화하는데 강점이 있으며 통계 분석을 위해서 SAS 또는 SPSS 패키지를 직접 이용할 수 있는 자료 구조로 변환하는 S/W가 설치되어 있다. Blaise의 또 다른 장점은 다양한 분야의 통계 조사를 단일 통계 조사 시스템에서 실행할 수 있다는 점이며 가구 조사, 사업 및 경제 조사, 건강 조사, 노동력 조사, 에너지/환경/농업 조사와 다양한 자료 수집 기법을 BLAISE 시스템에서 실행할 수 있어서 대규모 조사 회사에서도 운용하고 있는 시스템이다.

BLAISE가 windows환경에서 작동될 수 있으나 프로그램을 직접 작성해야 하는 불

편한 점을 보완하고 컴퓨터 입력 방법도 음성 또는 문장을 직접 인식할 수 있는 입출력 기술의 발달을 수용할 수 있는 개념과 많은 선택 과정을 메뉴 방식에 의해서 운용할 수 있는 조사기획시스템을 개발해야 할 것이다.

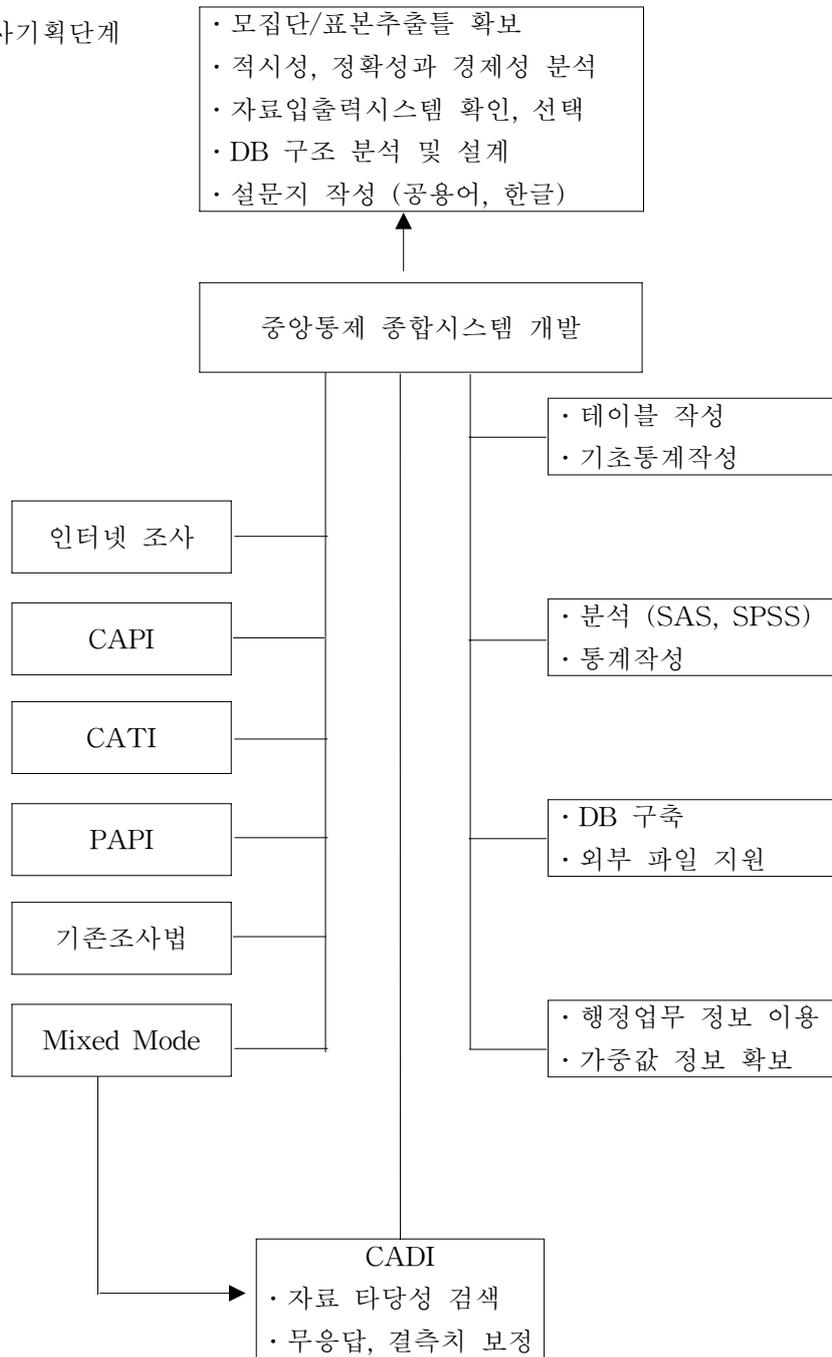
자료 수집 방법도 BLAISE 시스템에서 고려한 방법을 기본으로 하고 추가로 인터넷 조사를 포함하여 인터넷 시대에 실행 가능한 조사 기법을 포함해야 할 것이며, 추정 방법도 새로운 개념인 소지역 추정법이나 베이지안적 추정법을 적용할 수 있도록 외부의 행정 업무 자료와 인구 주택 총조사 등의 총괄적이고 기초적인 정보를 이용할 수 있는 알고리즘과 S/W가 설치되어야 할 것이다. 이상에서 설명한 컴퓨터 이용 통계 조사 종합 시스템의 개념도를 <그림 1>에 대략적으로 그려보았다.

<그림 1>은 크게 4단계로 나누어 표현하였다. 조사 기획 단계는 통계 전문가와 조사 업무 분야 전문가들의 활동을 표현하고, 상호보완적인 업무 협조가 필요하며, 자료 수집 단계는 필요 정보의 내용과 컴퓨터 기술의 현상태를 고려하여 자료 수집 방안을 결정하고 비표본 오차 관리에 유의해야 할 것이다. 자료 처리 단계는 표준 자료 입력 시스템, CADI와 CADE, 일괄 편집과 결측치 대체법 선택 등이 처리되고 자료 확인 검사를 완벽하게 처리해야 한다. 자료 분석 및 통계 생산 단계에서는 시각 효과를 높일 수 있는 테이블 작성과 SAS와 SPSS를 적용한 통계 분석 및 외부의 행정 업무 정보를 이용한 가중 추정값의 계산 가능성을 검토 분석한다.

인터넷 이용률이 60% 이상 되고 개인주의 의식의 팽창으로 조사 환경이 열악해지고 있는 때일수록 시간과 공간의 제한이 없는 인터넷 조사는 활용 전망이 크므로 인터넷 조사를 자료 기존의 자료수집기법의 대안으로 생각할 수 있다.

컴퓨터 이용 통계 조사 종합 시스템에 대한 연구는 많은 재원과 시간이 필요하기 때문에 시간을 지체하다가는 정보화 시대에서 통계 선진국으로 발전은 어려울 것이다. 차세대의 통계조사 종합 체계의 연구는 앞으로 5년 이내에 실행을 목표로 두고 통계청 등의 국가 기관이 개발 연구를 주도하고 민간 여론 조사 기관들도 공동협력체계를 구성하여 통계 조사 종합 시스템을 연구하거나 여론 조사 기관을 대형화하여 세계적으로 경쟁력 있는 통계 작성 기관으로 발전시켜야 할 것이다. 모집단 정의와 표본 추출률 확정에서 통계 작성 발표까지를 단일 종합 시스템에 의해서 수행할 수 있도록 전문 인력 양성 및 관리에도 관심을 가져야 할 것이다.

(1) 조사기획단계



2) 자료 수집 단계

3) 자료 처리 단계

4) 자료 분석 및 통계 생산

<그림 1> 통계조사 종합시스템

6. 결론

21세기 지식 정보화 사회에서 필요한 지식의 내용이 문제 해결의 방법을 아는 지식인 방법지를 핵심 요소로 하고 있으므로 정보화 시대에 필요한 통계도 이론적인 기반 위에 실용성이 강조된 실용 통계가 될 것이며 실용 통계의 구비 특성인 적시성, 정확성과 경제성을 충족하는 통계를 생산하기 위해서 기획, 자료 수집과 통계 작성 등에서 컴퓨터의 이용이 활성화되고 있다. 앞으로 지식 정보화 사회에서 필요한 실용 통계를 원활하게 생산하여 활용하는 방안 모색을 위해서 현재 우리나라의 컴퓨터 보급과 이용 현황을 살펴보았다.

우리나라 컴퓨터 보급은 정부의 국민 인터넷 PC 보급 정책 시행으로, 가정까지도 컴퓨터 보급률과 이용률이 급증하여 정보화 사회의 기반 조성은 미국 등 선진국에 뒤지지 않은 것으로 나타났다. 특히 인터넷 이용률은 폭발적으로 증가하여 사이버 공간 사회의 도래가 멀지 않았음을 예고하고 있다. 현시점에서 컴퓨터를 이용한 통계 조사 기법의 적용 현황은 미국, 캐나다와 유럽 등에 비해서 저조한 실태이며 통계청 등 정부 통계 작성 기관과 민간 조사 회사 모두가 컴퓨터를 이용한 조사 기법의 적용에 대한 적극적이고 심층적인 연구와 투자가 부족하여 통계 선진국 수준의 통계 발달을 기대하기 어려운 것이 사실이다. 그러나 인터넷의 빠른 보급률과 이용률의 증가를 잘 이용하여 인터넷 이용 통계 조사 기법을 집중적으로 연구 발전시킨다면 통계 후진성을 벗어날 수 있을 것이다.

앞으로는 다양한 분야에서 적시성, 정확성과 경제성을 갖춘 실용 통계의 요구가 많아질 것이므로 이를 원활하게 생산할 수 있는 컴퓨터 이용 통계 조사 종합 시스템을 제안하였으며 본 종합 시스템과 유사한 개념을 네덜란드 통계청에서 10여 년 간의 집중적인 연구 노력 끝에 BLAISE 시스템을 개발하여 여러 나라에서 활용되고 있으나 본고에서 제안한 통계 조사 종합 시스템은 인터넷 조사를 추가하고 행정 업무 정보 등 외부의 기초적이고 대규모의 정보를 연계하여 기획과 추정 단계에서 이용할 수 있는 종합 정보망의 활용을 보완하였다.

통계 수요 분야가 다양해지고 각양각색의 통계를 작성하기 위해서 신뢰성있는 통계를 생산할 수 있는 자격을 갖춘 통계인이 필요하지만 현실적으로 쉽지 않기 때문에 민간 조사 기관들도 대규모 조사 전문 회사로 통합 조정되거나 통계 작성의 전문적인 컨설팅만을 담당하는 통계 전문 지식 공유를 위한 공동 협력 체제를 갖추어야만 적시성과 신뢰성을 갖춘 통계를 활용하여 지식 정보화 사회를 성공적으로 개척할 수 있을 것이다.

실용 통계를 원활하게 생산하고 어느 분야에서 어떤 종류의 통계라도 신뢰하면서 통계를 활용하기 위해서 통계 작성의 외형적인 기반 조성도 중요하지만 보다 더 중요한 것은 통계인들의 의식이 변해야 하고 또한 일반인들도 통계를 필수 불가결한 정보로 이용해야 한다는 의식이 있어야만 건전하고 신뢰성있는 통계를 생산하고 활용할 수 있을 것이다. 이제 통계는 형식적인 보조 정보가 아니고 의사 결정이나 정책 수행

의 성패를 좌우하는 핵심적인 인프라의 기능을 할 수 있도록 모든 통계인들이 통계의 전문성과 신뢰성을 높이는데 최선의 노력을 다해야 할 것이다.

참고문헌

- 김효근 (1998), 《신지식인》, 매일경제신문사
- 이계오 (2000), “통계 조사의 현황과 전망”, 《조사연구》, 1(1):1-20
- 이계오 (2000), “인터넷 여론조사의 현황과 전망”, 《Internet Survey Workshop 논문집》, 1-17
- 통계청 (1997), 《컴퓨터에 의한 현지 조사 개선 방안》
- 통계청 (2000), 《1998년 기준 종합 환경 지수 관련 자료》
- 한국정보문화센터 (2000), 《2000 국민 생활 정보화 현황》
- Comley, Pete (2000), "The Use of the Internet for Opinion Polls", <http://www.vitalsurveys.com>
- Couper, Mick P. (1998), "The History and Development of Computer Assisted Survey Information Collection Methods", 《Computer Assisted Survey Information Collection》, John Wiley & Sons, New York
- Westat (2000), "Blaise for Windows", <http://www.westat.com/blaise/>