

정기통계품질진단 연구용역

『민간기업정보보호실태조사』
2010년 정기통계품질진단
연구용역 최종결과보고서

2010. 10.

주 의

1. 이 보고서는 통계청에서 수행한 정기통계품질진단 연구
용역사업 결과보고서입니다.
2. 이 보고서에 대한 저작권 일체와 2차적 저작물 또는
편집저작물의 작성권은 통계청이 소유하며, 통계청은 정책상
필요시 보고서의 내용을 보완 또는 수정할 수 있습니다.

제출문

제 출 문

통계청장 귀하

본 보고서를 “민간기업정보보호실태조사 2010년 정기통계품질진단” 연구용역 과제의 최종 연구결과물로 제출합니다.

2010년 10월 19일

한국통계학회장 최 종 후 ㉠

연구진

책임연구원	신기일 (한국외국어대학교 통계학과)
연구원	정석오(한국외국어대학교 통계학과) 강석훈(성신여자대학교 경제학과) 오완근(한국외국어대학교 경제학과)
연구원 (표본설계진단)	이상은(경기대학교 응용정보통계학과)
연구보조원	박기덕(한국외국어대학교 통계학과) 김지학(한국외국어대학교 통계학과) 김민영(한국외국어대학교 통계학과) 김부식(한국외국어대학교 경제학과) 황지애(성신여자대학교 경제학과) 강대완(경기대학교 응용정보통계학과)

품질보고서

『민간기업정보보호실태조사』 품질보고서

2010. 10.

통 계 청



차 례

- 1. 개요 2
- 2. 통계품질정보 2
 - 가. 차원별 품질 상태 2
 - (1) 관련성 2
 - (2) 정확성 4
 - (3) 시의성/정시성 5
 - (4) 비교성 5
 - (5) 일관성 5
 - (6) 접근성/명확성 6
 - 나. 기타 품질관련 정보 6
- 3. 결론 6

1. 개 요

'민간기업정보보호실태조사'는 국내 민간 기업의 정보보호 정책 및 조직 구성 현황, 역기능 대응 실태와 정보화 역기능 피해현황 등을 종합적으로 파악하여 국내 기업의 정보보호 수준 및 정보보호 관련 정책 수립의 기초 자료를 제공하는 것을 목적으로 실시된다. 2009년 조사의 경우 전국의 종사자수 5인 이상 네트워크로 연결된 PC를 한 대 이상 보유한 사업체 308,529개 중 2,234개를 대상으로 방문면접조사(한국갤럽에 의뢰)가 2009. 11. 10 ~ 2009. 12. 30의 기간 동안 이루어졌다. 표본 추출은 다단계층화계통추출법을 기본으로 하였는데, 업종별-규모별로 2단 층화한 후 각 사업체들을 지역별로 정렬하여 계통추출하였다. 2009년 조사 결과를 담은 연구보고서는 <2009 정보보호 실태조사 - 기업편>으로 2010년 8월 현재 출판 중이다. 조사전담기관은 한국인터넷진흥원(KISA, 문의: 조사분석팀 02-405-6541)이며, 모든 공표 자료들은 한국인터넷진흥원홈페이지 <http://www.kisa.or.kr/> 에서 자유로이 내려받을 수 있다.

본 품질보고서의 목적은 관련성, 정확성, 시의성/정시성, 비교성, 일관성, 접근성/명확성 의 6개 차원에서 '민간기업정보보호실태조사'의 품질진단을 실시하여, 진단 과정에서 도출된 각종 자료, 전문 이용자 및 일반 이용자의 의견 등을 종합 분석하여 이용자들을 위해 통계의 질적 품질 정보를 제시하는 것이다.

2. 통계품질정보

가. 차원별 품질 상태

(1) 관련성

관련성(relevance)은 통계의 포괄범위와 개념, 내용 등이 이용자의 요구사항에 부합되는 정도를 말한다. 관련성 부문의 진단을 위해 주요 이용자 파악

여부, 이용자 요구사항 파악 여부, 공표 자료가 작성 목적 및 이용자 요구에 부합하는 지 등의 측면에서 살펴본 결과는 다음과 같다.

‘민간기업정보보호실태조사’의 주요 이용자는 정보보호 관련 정책 수립에 관계된 전문 이용자가 대부분이며, 작성 전담 기관인 KISA에서 양적으로 충분한 이용자 명부를 확보하고 있는 것으로 파악되었다. 500부 정도의 인쇄된 형태의 연구보고서를 주요 이용자들이 소속된 기관 및 관련 기관에 배포하고 있으며, 전자 간행물 형태로 KISA 홈페이지를 통해 자유롭게 내려받을 수 있도록 하였다. 진단 대상 통계의 연구 보고서를 포함해 관련 간행물들의 수준이 전반적으로 매우 높으며 통계의 작성 목적 및 이용자의 요구 수준을 대체로 만족시키고 있는 것으로 판단된다.

다만 특정 자료의 다운로드 건수 등 자료별 활용도를 파악할 수 있는 수치를 홈페이지 상에서 알 수 없기 때문에 자료별 활용도를 파악하기 어려웠던 점이 아쉽다. 또한 이용자 명부의 이용자 중 상당수가 허수인 것으로 드러났다. 작성기관에서 적극적인 이용자 층을 구별하여 별도의 명부를 관리하고 이들과 피드백을 주고받을 수 있는 절차적 체계를 갖추어 이용자 측면에서의 통계품질을 지속적으로 개선시켜나갈 기반을 마련할 필요가 있다.

그리고 통계표 상에 ‘사례수’로 표시된 수치들이 표본 조사 결과를 전체 모집단 크기로 투영한 ‘추정치’이다. 따라서 통계표 작성 시 ‘사례수’를 적절한 다른 용어로 대체할 필요가 있다. 끝으로, 조사표에서 조사된 내용의 상당 부분이 공표 자료에서 빠져 있는 점은 자료의 특성 상 발생 가능한 문제들을 고려할 때 이해가 되는 바이나 가능한 범위에서 최대한 공표 자료에 포함시킬 수 있는 방안을 마련할 필요가 있다.

이상을 종합할 때 ‘민간기업정보보호실태조사’의 관련성 차원에서의 품질은 상당한 개선의 여지가 있는 것으로 평가된다.

(2) 정확성

정확성이란 추정값과 미지의 참값에 근접하는 정도를 말한다. '민간기업정보호실태조사'의 경우 체계적 절차에 의해 정확성이 높은 통계를 생산하기 위한 노력이 이루어진 것으로 판단된다. 예컨대 사후층화를 통해 편차를 줄일 수 있도록 가중치를 조정하는 작업이 적절한 절차에 의해 체계적으로 이루어진 점은 높이 평가된다. 다만 다음의 세 가지 문제점이 지적된다.

첫째, 표본틀(sampling frame)이 두 단계로 되어 있어 이용자에게 혼란을 줄 수 있는 점이다. 연구보고서에 의하면 1차 표본틀은 '사업체기초통계조사(통계청)' 대상 사업체이고, 2차 표본틀은 '정보화통계조사(한국정보화진흥원)' 대상 사업체 중 네트워크 구축 사업체로 되어 있다. 그러나 2차 표본틀이 되는 '정보화통계조사'가 조사 대상 리스트를 제공하는 것이 아니라는 점과, 모집단 크기인 308,529개라는 수치도 추정치이기 때문에 이에 기초해 산정된 가중치의 정확도는 떨어질 수 밖에 없다는 것이 문제점으로 지적된다. 현재 통계 작성 여건 하에서 불가피한 조처임은 이해하지만, 이용자의 혼란을 최소화하려면 이 부분에 대한 설명을 공표 자료에 강조할 필요가 있다.

둘째, 표본의 크기가 너무 작다. 모집단 크기가 30만이 넘는데 2,500개의 사업체를 최종 표본으로 하여 추출률이 0.83%인데 이는 너무 낮은 수치이다. 이렇게 되면 범주별 표본오차를 제대로 관리하는 것이 불가능하며, 조사 내용의 성격 상 조사 시 응답의 신뢰도가 상당히 떨어질 수 밖에 없는 점을 등을 고려한다면 현행 표본 크기는 상당히 문제가 있다.

셋째, 공표 자료에 주요 변수별 상대표본오차를 제시하고 있으나 통계 수치들이 0과 1 사이의 값을 갖는 비율(proportion)이기 때문에 현행 공표 수준에서는 불필요하다.

이상을 종합할 때 정확성 차원에서의 품질은 상당한 개선이 요구되는 것으로

로 판단된다.

(3) 시의성/정시성

시의성은 작성기준시점과 결과공표시점 간의 시간경과 정도를 나타내는 개념이고, 정시성은 사전에 예고된 공표 날짜를 정확히 준수하는 가에 대한 개념이다. '민간기업정보보호실태조사'의 경우 시의성 및 정시성 차원의 품질에 전반적으로 큰 문제가 없으나, 공표일을 예고하고 있지 않아 공표 시기가 해마다 상당히 달라지는 점이 개선 사항으로 지적된다.

(4) 비교성

비교성은 시공간적으로 자료가 비교 가능한 정도를 말한다. '민간기업정보보호실태조사'는 해당 분야에서 국내외적으로 거의 유일한 통계이다. 따라서 공간(지리)적 비교성을 따지는 것은 현재로서는 큰 의미가 없다. 다만 시간적 비교성 측면은 통계 생산기간이 길지 않아 평가 항목이 안정화되지 못한 점을 감안하더라도 조사표 내 주요 항목의 응답 형식이 해가 바뀌면서 달라지는 경우가 있어 시계열적 정보를 얻는 데 어려움이 있다. 주요 항목에 대한 응답 형식을 조속히 확정하고 자료가 축적됨에 따라 향후 시계열적 분석이 가능하도록 데이터베이스를 잘 관리해 나갈 필요가 있다.

(5) 일관성

일관성은 서로 다른 기초자료나 작성방법에 따라 작성된 통계자료라 하더라도 동일한 사회 현상을 반영하는 경우 각 통계 자료가 서로 얼마나 유사한지를 나타내는 정도를 말한다. 비교성 진단에서 지적된 바와 같이 비교 가능한 통계가 거의 없기 때문에 일관성 차원에서의 품질 평가는 별 의미가 없는 것으로 판단된다. 다만 국가 승인 통계로서 일반적으로 갖추어야 하는 작성 주기 및 산업 분류 등의 통계 작성 기준의 적정성, 통계 내 통계 수치 간 일치성 등을 사업체를 대상으로 하는 다른 통계들과 견주어 볼 때 손색이

없는 수준으로 관리되고 있다고 판단된다.

(6) 접근성/명확성

접근 가능성은 이용자가 데이터에 손쉽게 접근할 수 있는 정도를 말하며 명확성은 통계표와 메타정보 등 통계 자료의 이용 편의성 제공 정도를 나타내는 개념이다. '민간기업정보보호실태조사'의 경우 인쇄된 형태의 연구 보고서 배포를 포함해 각종 자료를 온라인 상에서 손쉽게 내려받을 수 있도록 되어 있어 접근성 및 명확성 차원에서 뛰어난 품질을 가진 것으로 평가된다.

나. 기타 품질관련 정보

'가'에서 다룬 6개 품질차원 외에 효율성(efficiency) 차원의 품질에 상당한 문제점이 있다. 작성기관 측이 확보하고 있는 우수한 품질관리기반과 상당한 수준의 조사 비용 등을 고려할 때 공표 자료의 정보량의 수준이 다소 미흡하다. 전수 조사 수준의 조사가 이루어지지 않는 한 공표수치의 신뢰도 확보가 거의 불가능한 본 통계의 특성을 고려할 때 불가피하다는 점은 인정되나, 조사표를 통해 조사된 항목 중 공표 시 누락되는 것들이 많고 그렇지 않은 항목들 역시 양적 정보를 질적 정보로 변환하여 정보량 손실이 상당한 상태로 공표되고 있는 점은 장기적으로 해결 방안을 찾아야 할 것으로 보인다.

3. 결론

'민간기업정보보호실태조사'의 차원별 품질을 진단한 결과 접근성/명확성은 우수한 것으로 판단된다. 비교성 및 일관성은 해당 분야에서 현재 생산되고 있는 국내외 공식 통계 중 거의 유일한 통계인 점을 고려할 때 깊이있는 진단이 어려운 것으로 보인다.

관련성은 자료별 활용도 관리가 미흡한 점, 공표 자료의 부적절한 용어 사용 등이 개선되어야 할 사항이다. 그리고 정확성의 경우 표본추출틀 문제, 표본

규모 문제 등이 중요한 개선점으로 도출되었다. 마지막으로 시의성/정시성은
공표 시기가 매년 유동적인 점이 개선사항으로 지적되었다.

요 약 문

최종결과보고서 요약문

연구과제명	「민간기업정보보호실태조사」 정기통계품질진단
주 제 어	정보보호, 표본 규모, 표본 설계, 정보량
연 구 기 간	2010년 4월-2010년 10월
연 구 기 관	한국통계학회
연구진구성	책임연구원 : 신기일(한국외국어대학교 통계학과) 연구원: 정석오(한국외국어대학교 통계학과) 연구원: 강석훈(성신여대 경제학과) 연구원: 오완근(한국외국어대학교 통계학과) 연구원: 이상은(표본설계담당, 경기대학교 응용정보통계학과)
<p>민간기업정보보호실태조사는 국내 민간기업의 정보보호 수준 및 정보보호 관련 정책 수립의 기초 자료를 제공하는 통계청 승인통계이다. 따라서 본 통계는 우리 나라 정보보호 수준을 전체적으로 파악할 수 있는 매우 중요하고 활용도가 높은 통계이므로 다양한 품질 차원 및 절차적 측면에서 체계적으로 점검할 필요가 있다. 이를 위해 통계청이 제시한 정기통계품질진단 핸드북을 바탕으로 통계품질진단을 실시했다.</p> <p>진단 결과를 정리하면 본 통계는 전반적으로 우수한 품질을 유지하고 있는 것으로 평가된다. 다만, 진단 과정에서 전국 단위 조사임에 비해 표본 규모가 작은 문제, 공표 자료의 정보량이 충분하지 않은 점, 표본 설계에서 개선이 필요한 일부 문제점 등이 발견되었다. 표본 규모를 적정한 수준으로 확보해 통계의 정확성을 확보함과 동시에 정보량이 풍부한 통계를 생산하고, 이용자 계층을 확대해 통계의 활용도를 제고하는 한편, 국제적 표준을 선도할 수 있는 품질을 확보할 수 있도록 지속적인 노력이 필요하다.</p>	

차 례

제 1 장 개 요	1
제 1 절 품질진단 개요	1
제 2 절 진단 대상 통계 개요	3
제 3 절 중점 진단사항	4
제 2 장 품질진단 결과	6
제 1 절 진단 결과 종합	6
제 2 절 통계품질 개선 과제	42
제 3 절 통계품질 개선 방안	46
제 3 장 발전 전략 및 로드맵	50
제 1 절 발전 전략	50
제 2 절 로드맵	53
참고문헌	55
부록 1. FGI 결과보고 (전문이용자)	56
부록 2. FGI 결과보고 (일반이용자)	61
부록 3. 수집자료 정확성진단 점검표	66
부록 4. 표본설계 진단 결과	71

표 차례

<표 1> 품질진단 방법 개요	1
<표 2> FGI 이슈별 논의 사항	13
<표 3> 이용자 만족도 조사 항목	14
<표 4> 이용실태 분석	16
<표 5> 종합만족도 산출 결과	18
<표 6> 종합만족도 및 전반적(체감) 만족도 비교표	18
<표 7> 세부작성절차별 점검 지표 및 점수	24
<표 8> 작성절차별 품질진단	27
<표 9> 품질차원별 품질진단	28
<표 10> 진단대상 통계 개요	29
<표 11> '정보화 투자 대비 정보보호 투자 비율'항목의 표본오차(RSE)	31
<표 12> 조사 전·후의 표본 층별 분포 비교	38
<표 13> 통계품질 개선과제	49

그림 차례

<그림 1> 항목별 중요도 비교	17
<그림 2> 포트폴리오 행렬	20
<그림 3> 공표범위 세분화 필요성	21
<그림 4> 원자료 제공 필요성	21
<그림 5> 재이용 의사	21
<그림 6> 작성절차별 품질진단	27
<그림 7> 품질차원별 품질진단	28
<그림 8> 목표 개념	50
<그림 9> 5개 발전과제와 목표 개념	52
<그림 10> 5개 발전과제 및 로드맵	54

제 1 장 개 요

제 1 절 품질진단 개요

통계품질은 관련성, 정확성, 시의성/정시성, 비교성, 일관성, 접근성/명확성의 여섯 차원을 고려해 진단하게 된다. 이를 위해 통계청에서는 품질관리기반, 이용자 만족도 및 요구사항 반영 실태, 세부 작성절차별 체계, 수집 자료의 정확성, 통계자료 서비스 등 다섯 부문에 대한 품질진단방법을 제시하고 있다. 여기에 통계의 정확성에 직접적으로 영향을 주는 표본 설계 부분 진단을 추가하여 진단을 실시하였다. (<표 1> 참조)

<표 1> 품질진단 방법 개요

품질관리 기반	- 통계작성현황 및 조직관리 실태 파악 - 인적·물적 자원 확보 현황 파악 - 애로 사항 파악	<ul style="list-style-type: none"> ☞ 품질관리기반 현황표 활용 (작성여건, 담당자 인식조사) ☞ 현장 방문 및 담당자 면담
이용자만족도 및 요구사항 반영 실태	- 통계이용 실태 및 요구 사항 파악 - 이용자 만족도 측정 및 의견 수렴 - 이용자 명부 작성 및 검토 - FGI를 통한 전문 및 일반 이용자 요구 사항 파악	<ul style="list-style-type: none"> ☞ 품질진단서 활용 ☞ 이용자 만족도 조사 실시 ☞ FGI를 통해 해당 통계의 문제점과 개선 아이디어 수집.
세부작성 절차별 체계	- 통계작성 기획 및 설계 - 자료수집 대상 및 기준 설정 - 자료입력 및 처리 - 통계자료 공표, 자료제공 및 사후관리	<ul style="list-style-type: none"> ☞ 품질진단서 활용 (조사통계용/가공통계용) ☞ 내부, 외부, 연구진 진단
수집 자료의 정확성	- 비표본오차 및 오류 점검 - 자료수집시스템의 적합성 진단	<ul style="list-style-type: none"> ☞ 품질 진단 양식 및 수정된 진단 양식 활용 ☞ 현장 방문 및 담당자 면담 ☞ 발생원인별 비표본오차 점검
통계자료 서비스	- 통계자료 수치오류 점검 - 이용자 정보제공 정도 점검	<ul style="list-style-type: none"> ☞ 발간물오류점검 지침 활용 ☞ 발간물 점검표 활용
표본 설계 진단	- 통계 작성 목적에 부합된 표본 설계 - 표본의 대표성과 상대표준오차 점검 - 적정 표본 수	<ul style="list-style-type: none"> ☞ 품질진단서 활용 ☞ 원자료 분석을 통한 최적의 표본 설계 진단
종합 진단 및 통계별 품질 개선 전략 제시		

품질진단은 그 과정에서 연구진과 진단대상 통계 작성기관과 협력하여 함께 통계 품질의 문제점과 개선 방안 등을 마련함으로써, 진단 후 관련 통계의 품질이 획기적으로 개선되고 통계 작성기관의 사기도 높아질 수 있도록 하는 동반자 입장의 선순환적 진단이 되어야 한다. 또한 품질진단 결과를 바탕으로 고품질의 국가통계작성 시스템을 어떻게 구축해야 하는지에 대한 의견을 제시함으로써 향후 신뢰성 높은 국가통계 품질관리시스템을 구축하는데에도 기여하도록 해야 한다. 통계 품질 진단 관련 자료는 문서화되고 모든 이해 관계자들에게 공유되어 유사 통계 작성 예정자가 활용할 수 있도록 해야 한다. 통계작성자가 통계작성 과정의 강점과 약점을 파악토록 하여 약점으로 진단된 부분을 적극적으로 개선토록 유도함으로써 해당 통계 품질의 중요성을 인지하도록 해야 한다. 또한 통계 이용자가 통계 작성과정을 이해함으로써 보다 높은 수준의 활용 및 분석이 가능한 점을 인식시키고, 우수한 이용자가 존재해야만 해당 통계가 지속적으로 발전할 수 있다.

통계품질은 생산자 측면, 이용자 측면에서의 품질이 모두 중요하다. 따라서 해당 통계의 이용 빈도가 높은 정보통신 전문가와 함께 통계 생산 과정을 진단하기 위한 통계 전문가를 모두 포함하여 연구진을 구성함으로써, 보다 종합적이고 다면적인 품질진단이 가능하도록 하였으며 생산자와 이용자 모두에서 발생할 수 있는 문제 해결이 가능하도록 하였다. 또한 이용자 자문 집단을 운영하여 이용자 측면의 통계 품질 향상을 도모하며 표본 설계 전문가를 포함한 통계전문가가 심층 진단하도록 했다.

본 진단 결과를 통해 통계작성기관이 스스로 품질진단 결과를 종합·분석함으로써 해당 통계의 품질 현황을 파악하고 품질을 개선시킬 수 있는 방안을 마련할 수 있는 기반을 마련하여 '민간기업정보보호실태조사'의 전반적 개선을 도모할 수 있을 것으로 기대된다. 또한 통계 품질진단과정을 통해 통계작성기관이 통계품질의 중요성에 대해 인식을 공유하고 통계 작성자로서의 자긍심도 고취하게 될 것이다. '민간기업정보보호실태조사'에 대한 체계적인 품질진단 시스템을 구축을 통해 국가통계 품질평가가 체계적으로 이루어지도록

록 노력하며 향후 정보통신분야에 적합한 품질진단체계를 마련하게 될 것으로 기대된다.

제 2 절 진단 대상 통계 개요

1. 조사 목적

‘민간기업정보보호실태조사’는 국내 민간기업의 정보보호 정책 및 조직 구성 현황, 역기능 대응 실태와 정보화 역기능 피해 현황 등을 종합적으로 파악하여 국내 기업의 정보보호 수준 측정 및 정보보호 관련 정책 수립의 기초자료를 제공하는 것을 목적으로 한다.

2. 조사 대상 및 조사 내용

2009 조사의 경우 전국의 종사자 수 5인 이상, 네트워크로 연결된 PC를 한 대 이상 보유한 사업체 2,234개를 대상으로, 2009.11.10.~2009.12.30.의 기간에 방문면접조사로 조사가 이루어졌다. 주요 조사 내용은 다음과 같다.

- 정보보호 정책 수립 및 정보보호 조직 구성 현황
- 정보보안 환경 평가 및 위협 요소 평가
- 임직원 대상 정보보호 교육 실시 및 정보보호에 대한 투자 현황
- 정보보호 제품 사용 현황 및 정보보호 업무 수행 방식
- 보안패치 적용 방식 및 정보시스템 사용자 인증 기법
- 웹사이트를 통하여 수집한 개인정보의 유/노출 방지를 위한 대응 현황
- 인터넷 침해사고 대응을 위한 활동 및 복구계획 수립/운영 현황
- 인터넷 침해사고 및 개인정보 유·노출 피해 경험

3. 작성 및 공표 체계

조사 기획은 방송통신위원회와 한국인터넷진흥원(이하 KISA) 조사분석팀이 담당하고 있고, 현장 조사와 자료 처리는 외부 용역을 주고 있는데 2009년 조사는 한국 갤럽이 담당하였다. 결과 분석 및 공표는 방송통신위원회와 KISA 조사분석팀이 공동으로 담당한다. 최종연구보고서인 <정보보호실태조사 - 기업편> 이 인쇄간행물 및 한국인터넷진흥원 홈페이지의 자료실을 통해 공표되고 있다(<http://www.kisa.or.kr>).

제 3 절 중점 진단사항

1. 표본설계 부문

○ 목표모집단과 조사모집단: 표본 추출틀로 사용되는 '사업체기초통계조사'와 '정보화통계조사'의 표본틀 작성 과정 및 중복성 여부와 모집단의 타당성 및 시의성을 검토한다.

○ 표본설계: 표본설계의 타당성 검토를 위해 사후층화방법의 타당성과 허용오차의 수준을 검토한다.

○ 조사방법: 조사원 교육내용 및 업무량을 점검한다. 설계 당시 표본과 실사 표본 매칭 비율 및 대체 타당성 검토한다. 조사 오류의 체크 시스템을 확인하고 사후 조사를 통한 정확성 확인한다.

○ 추정방식: 표본설계와 부합되는 추정방식을 적용했는지의 여부와 표본 오차 공식을 검토한다.

○ 표본의 크기: 공표 단위를 기준으로 표본오차/상대표준오차를 산출하여 표본규모의 타당성을 검토한다.

○ 무응답처리: 전수조사 층에서 무응답이 발생한 경우 적절한 방식으로 처리되었는 지 검토한다.

2. 표본 설계 외 부문

○ 시계열 자료: 민간기업정보보호실태조사의 경우에는 횡단면적으로 현 상황을 파악하는 것과 동시에 시계열적으로 어떠한 변화가 있는 지를 파악하는 것이 매우 중요하다. 그러나 본 통계의 경우 연도별로 설문내용 등에서 변경이 많아 시계열적 분석이 어려운 경우가 다수 존재한다. 예를 들어 '정보보호 정책 및 지침 수립 여부'는 가장 기본적인 항목으로 판단되는데, 2008년 보고서의 경우 응답기준이 변경되어 연도별 시계열을 주의하라는 설명과 함께 당해연도의 결과만을 제시하고 있다. 주요 내용 설명 뒤에 나오는 통계표 역시 거의 모든 항목이 당해연도 자료만 제시되고 있어 시계열분석이 불가능한 구조이다. 다만 주요 내용 소개에서 일부 항목에 대해 승인 시점인 2007년도 자료부터 3년간 자료가 제시되고 있다

○ 기업 특성 및 조사 목적을 반영한 조사: 보다 기업의 특성을 많이 조사할 필요가 있다. 현재 본 통계에서는 업종과 규모를 기준으로 하고 있으나, 본 통계 관련 정보를 활용해 조사목적에 맞는 조사 재설계가 필요할 것으로 보인다. 예를 들어 정보보호가 필요한 개인정보가 10만명 이상인 기업과 이하인 기업으로 분류하여 전자의 경우에는 보다 자세한 내용을 조사하고 후자의 경우에는 조사내용을 간략히 하는 등의 방식으로 조사구조를 검토할 필요가 있다.

제 2 장 품질진단 결과

제 1 절 진단 결과 종합

1. 품질관리기반 진단

1.1 진단 개요

‘민간기업정보보호실태조사’의 품질관리기반 진단을 위해 중점적으로 살펴본 사항은 다음과 같다. 먼저 통계에 대한 기관장의 관심정도를 파악하고 기관장이 통계의 목적과 그 중요성을 인식하는지를 파악하였다. 즉 통계담당 직원들이 고품질의 통계를 작성하도록 업무 환경을 조성해 주는지를 직간접적으로 파악하고, 적절한 수의 통계 인력이 확보되었는지, 통계 관련 분야 전공자 및 전문가를 확보하여 통계작성 업무에 투입하고 정기적으로 적절한 교육 및 훈련을 실시하고 있는지를 살펴보았다.

위 과정은 통계청에서 제시한 품질관리기반 현황표(①기본현황, ②통계작성 여건, ③조직관리실태 및 통계작성 담당자 인식, ④통계작성 담당자와의 면담 내용)를 바탕으로 이루어졌다. 먼저 통계 생산기관 담당자가 작성을 하고 연구진이 면담을 통해 통계 작성 시 애로사항과 품질관리기반 확보를 위한 개선사항을 추가로 작성하였다.

1.2 현황 및 문제점

가. 인적여건

KISA에서는 5명의 직원이 기획과 행정을 맡고 있다. 통계작성 인력의 통계 업무 전담 정도는 전체 업무의 약 36%이며 통계관련 업무 경력은 평균 4.4년 정도였다. 작성팀에 통계학 전공자가 포함되어 있어 물적 자원을 효율적

으로 활용하고 있는 것으로 파악되었다. 인적 여건의 양적 수준과 질적 수준이 모두 양호한 것으로 판단된다.

나. 물적 여건

통계작성관련 예산의 대부분이 외부기관위탁에 사용되고 있다. 예산 규모는 매년 증가 추세에 있다. 또한 정보자원을 살펴보면 상당한 수준의 업무 시스템, 통계처리를 위한 전문 패키지, 양호한 스펙의 운영 장비를 확보하고 있다. 예산 및 정보자원 확보 면에서 볼 때 통계 작성을 위한 물적 여건은 양호한 것으로 판단된다.

다. 조직관리 실태 및 통계작성 담당자의 인식

대체로 조직의 리더와 담당자 모두 진단 대상 통계의 중요성에 대해 인지하고 있으며, 통계 업무를 원활히 수행할 수 있는 환경을 갖추고 있는 것으로 평가된다. 품질진단을 위한 연구진의 각종 요구와 질의에 대해 매우 적극적이고 효과적으로 응대하였으며, 이 과정에서 조직 및 업무 환경의 관리가 잘 이루어지고 있음이 확인되었다.

라. 애로사항

급변하는 IT 환경과 밀접하게 관련되어 있는 본 통계의 특성 상 변동 사항이 빈번하게 발생하는 점을 고려할 필요가 있다. 변동 사항이 발생할 경우 매번 승인 변경 신청 프로세스를 거쳐 통계청과 협의해야 하는데, 변동 내역을 따져 통계청의 협의가 필요한 사항과 작성기관의 자율성이 확보되는 사항을 구분할 필요성이 있다는 작성 기관 측의 건의가 있었다.

2. 이용자 만족도 및 요구사항 반영 실태

2.1 진단 개요

이용자의 통계에 대한 만족도 및 요구사항 반영 정도를 측정하기 위하여 전문 이용자 및 일반 이용자를 대상으로 표적집단면접(focus group interviews; 이하 FGI)과, 확보된 이용자 명부를 활용해 통계 이용자를 대상으로 하는 전반적인 만족도 조사를 실시했다. 조사에 참여한 이용자들은 대체로 본 통계를 적극적으로 이용하고 있는 전문이용자가 대부분이었기 때문에 상당한 성과를 거둘 수 있었다.

일반이용자 FGI는 2010년 6월 1일 14:00~15:15 한국인터넷진흥원 서초청사 회의실에서 열렸다. 참석자는 학계 3명, 관련 기관 연구원 3명, 통계청 담당자 1명 등 모두 7명으로 구성되었다. 전문이용자를 대상으로 한 FGI는 2010년 5월 18일 14:20~16:50 한국인터넷진흥원 서초청사 회의실에서 열렸다. 참석자는 모두 10명으로 학계 4명, 관련기관 연구원 3명 등으로 구성되었다.

이용자만족도 조사를 위해 최초 제공받은 이용자 명부는 32명이었다. 이후 이용자 만족도 조사과정에서 최소 60명의 응답자를 확보하기 위해 관련 기관 및 회사의 관계자들과 FGI 참석자 및 외부진단위원 등 68명의 명부를 추가로 확보하여 총 100명의 이용자 명부가 작성되었다. 이를 조사전문업체에 전달하여 이용자 만족도 조사를 실시한 결과 100명 중 무응답자가 40명이었다. 무응답 요인은 '이용 경험이 없는 경우'가 32건, '바빠서'인 경우가 6건, '장기 출장'인 경우가 2건이었다. 무응답 요인 중 가장 많은 비중을 차지하는 '이용 경험이 없는 경우' 즉 실제 이용자가 아닌 경우는 최초 이용자 명부 32명 중에서 15명(47%), 추가된 68명 이용자 명부에서 17명(25%)이었다. 최초 이용자 명부의 절반에 가까운 47%가 실제 이용자가 아니었다는 점은 문제가 있다. 작성 기관이 이용자 그룹을 효과적으로 파악하고 있지 못하고 있는 점을 방증하기 때문이다.

2.2 현황 및 문제점

가. 표적집단면접

일반이용자 FGI와 전문이용자 FGI로 구별해 2차에 걸쳐 실시한 FGI에서 논의된 주요 이슈를 정리하면 다음과 같다.

첫째, 모집단 및 표본추출틀의 문제이다. 사업체기초통계와 정보화통계조사 의 두 개의 표본추출틀을 병용하는 것은 주어진 여건에서 최선임은 인정되나, 이용자 입장에서 자료 이용 및 해석 시 혼란이 생길 가능성이 있다는 점이 지적되었다. 가능하다면 '민간기업정보호실태조사'가 목표로 하는 모집단의 리스트 제공이 가능한 기초통계를 확보할 필요와 그 실현 가능성에 대한 논의가 이루어졌다. 또한 '기업'과 '사업체'의 개념은 엄연히 구별되는 것임에도 불구하고, 조사는 사업체를 대상으로 실시하면서 통계명에 '민간기업...'이라는 이름을 사용하는 문제가 지적되었다. 신뢰할 만한 기업 리스트를 확보할 수 있다면 해결 가능할 것이나 아직 우리 나라 공식 통계로는 존재하지 않다는 한계가 지적되었다. 무엇보다 조사 대상이 바뀌면 그 간의 통계와 단절이 생겨 현실적으로 개선의 여지가 별로 없는 것으로 판단된다.

둘째, 조사표 설계 및 조사 항목의 문제이다. 조사표에서 조사하는 내용이 상당히 많고 민감한 내용들로 구성되어 있어 피조사자 입장에서 굉장히 귀찮을 것이기 때문에 효과적인 데이터가 나올 수 없을 거라는 우려가 제기되었다. 웹상에서 직접 확인 가능한 항목들은 조사표 문항에서 삭제하거나, 조사 항목을 구별해 일부 항목은 격년으로 조사 주기를 늘림으로써 피조사자의 부담과 조사 비용을 함께 줄일 수 있는 방안을 연구해 볼 만하다는 의견이 제시되었다. 사업체 규모에 따라 의미가 없는 항목들이 조사 항목에 포함되어 있다는 점이 지적되었다. 예를 들어 CSO제도 관련 항목들을 100명 이하 규모 업체를 대상으로 조사하는 것은 무의미하다는 것이다. 이러한 항목

들은 조사표 설계 시 사업체 규모에 따라 따로 응답하도록 하는 것이 효과적이라는 의견이 제시되었다. 규제 기관 측에서 조사를 하기 때문에 피조사자 입장에서 민감한 항목에 대해 방어적으로 응답을 하게 되는 문제가 있음이 지적되었다. 이를 고려하여 방어적 응답을 사전에 차단할 수 있는 장치를 조사표 설계 시 도입할 필요가 있다는 의견이 제시되었다. 아울러 조사표 구성 단계에 민간기업의 정보보호팀의 관계자들이 직접 참여시킴으로써 현실적 활용도가 높은 통계로 발전시키는 방안이 논의되었다. 업체 내부 종사자에 의한 사고나 지능적인 범죄 등을 예방하기 위해 미리 엄중한 감사 방법을 설정해 두는 내부 통제(internal control) 관련 항목이 조사 항목에 추가될 필요도 제기되었다.

셋째, 표본 규모의 문제이다. 전국 조사임에 비해 표본 크기가 작은 것이 아쉬운 점으로 지적되었다. 국내외 거의 유일한 통계이고 국가적으로 활용도가 높음을 고려해, 예산을 추가 확보하고 예산 집행을 효율화하여 보다 정교하고 효과적인 정보를 제공하는 통계로 발전시켜 나갈 필요가 있다는 의견이 제시되었다. 구체적 규모는 표본오차를 고려해 표본규모를 재설정할 필요가 있음이 지적되었다. 단기적으로 표본 크기를 늘리는 것이 현실적으로 불가능하다면 패널(panel) 조사 등을 통해 비표본오차를 줄이고 관련성 부문의 품질을 제고할 수 있다는 의견이 제시되었다.

넷째, 공표 자료의 정보량 보강의 문제이다. 사업체규모별보다는 고객정보를 많이 가지고 있는 업체냐 아니냐, 혹은 IT 의존도가 어느 정도인지에 따라 자료가 제공되는 것이 이용자 입장에서 더 의미가 있다는 의견이 있었다. 그러나 표본 설계 단계에서 이를 고려하기 어렵기 때문에 일부 산업에 한해 세분화된 자료를 제공하는 방안이 제시되었다. 즉 운수와 통신업을 묶어서 통계가 제공되고 있는데, 이 통계의 특성을 고려해 통신업을 분리해서 조사하고 통계 자료를 제공하는 아이디어가 논의되었다. 그리고 전체 업종에 대한 업종별×규모별 2차원 교차표(cross-table)는 현행 조사의 표본 규모를 고려할 때 불가능하지만 최소한 본 통계에 관련 업종에 한해서라도 규모별 정

보를 동시에 고려한 2차원 정보를 제공하는 것이 제안되었다. 보고서 내 통계표의 대부분의 변수가 범주형이고 단순 사례 수이거나 비율 형태의 수치만 제공되고 있는데, 양적(quantitative) 변수 형태의 자료도 제공될 필요가 있다는 점이 지적되었다. 아울러 정보보호 수준을 대표할 수 있는 지수(index)가 개발된다면 활용 가치가 매우 높을 것이라는 의견이 있었다. 공표 항목 중 주요 항목에 대해서는 연도별 추세 및 추세의 일관성을 파악할 수 있도록 시계열 형태의 자료도 함께 제공하는 것이 제안되었다. 공표자료를 온라인 상에서 pdf로만 제공하는 것보다는 HWP, Microsoft Word, Microsoft Excel 등으로도 제공해 이용자들이 편리하게 이용할 수 있도록 해야 한다는 의견이 제시되었다.

다섯째, 원자료 공개 문제이다. 통계법에 공개를 원칙으로 하고 있는 바에 따라 원자료의 공개에 대한 요구 사항이 나왔다. 다만, 개별 기업의 특정 정보가 포함될 수도 있고 정보보호 쪽 정보는 민감한 부분이 많기 때문에 상당한 어려움이 예상되었다. 그러나 원자료를 어느 수준까지 가공해 제공할 것인지에 대한 기술적 고민을 거친 후 원자료 제공의 원칙은 지키는 것이 바람직하다는 의견이 지배적이었다. 공표된 자료 중 이용자 입장에서 불필요한 내용이 많고 오히려 실제로 필요한 내용이 없는 부분도 있을 터인데, 원자료를 제공하게 되면 이용자 스스로 필요한 수치들을 계산해 사용할 수 있다는 이용자 쪽에서의 장점이 있으며, 아울러 해당 통계 활용도가 높아지게 될 것으로 예상되므로 작성기관에게도 유리한 점이 있음이 함께 지적되었다. 국가통계 DB 구축이 완료되면 이를 통해 원자료 제공이 가능하게 될 것이라는 전망도 제시되었다. 한편 동일 자료에 이용자들이 각기 다른 분석틀을 적용해 같은 문제에 대해 이용자마다 전혀 다른 수치를 가공해 낼 가능성이 매우 높는데 정보보호 관련 수치들이 갖는 민감한 성격 때문에 실무자 입장에서 매우 조심스럽다는 의견도 있었다.

여섯째, 부적절한 용어 사용 문제이다. 통계표에서 '사례수'라는 용어가 반복적으로 사용되었는데 세심한 검토 과정을 거치지 않고 단순히 통계 패키지

의 output을 그대로 복사해 붙여넣은 것이 아닌지 의심되고, 대부분 '사례수'가 아닌 조사결과를 모집단에 투영한 추정치인 경우가 많아 용어를 수정해야 한다는 지적이 있었다.

이상을 요약하면 <표 2>와 같다. 논의된 사항을 품질차원별로 구분해 볼 때 특기할 만한 것은, 6개 품질차원 중 관련성, 정확성, 접근성/명확성에 관한 내용이 대부분이었고, 나머지 차원인 시의성/정시성, 비교성, 일관성 등에 대해서는 거의 논의가 없었다는 점이다. 시의성/정시성 차원의 품질이 현재의 이용자들을 대체로 만족시키고 있고, 비교성 및 일관성 차원은 비교 가능한 여타 통계나 표준이 아직 없기 때문인 것으로 분석된다. 향후 이용자층이 다양해지게 되면 재차 진단이 필요할 것으로 판단된다.

<표 2> FGI 이슈별 논의 사항

이슈	논의 사항	해당 품질차원
모집단 및 표본추출틀	목표모집단이 일치하지 않는 ‘사업체기초통계조사’ 와 ‘정보화통계조사’ 를 1, 2차 표본추출틀로 사용	정확성
	통계명이 ‘민간기업...’ 이지만 조사 대상은 사업체임	정확성
조사표 설계 및 조사 항목	피조사자 부담 경감을 통해 수집 자료의 신뢰도 향상	관련성
	사업체 특성에 따라 무의미한 조사 항목이 상당 수임	관련성
	피조사자의 방어적 응답을 방지할 수 있는 장치 마련	정확성
	조사표 설계에 민간기업 정보보호 담당자들 참여 필요	관련성
	조사항목에 ‘내부통제’ 관련 항목 추가할 필요	관련성
표본 규모 및 비표본오차	전국 단위 조사임을 고려할 때 표본규모 작음	정확성
	주요 항목의 표본오차를 고려해 표본 규모 재설정 필요	정확성
	패널 조사를 통해 비표본오차를 줄이고 관련성 차원의 품질 제고	정확성 관련성
공표자료 정보량	사업체 규모별 정보보다는 고객정보를 많이 가진 정도, IT 의존도에 따른 정보가 의미있음	관련성
	운수/통신업을 묶어 통계를 제공하는 데 통신업을 따로 구별해서 제공할 필요	관련성
	유관 업종에 한해 다차원 교차 정보(표) 제공 필요	관련성
	피해액 등 양적 변수 자료 제공 필요	관련성
	정보보호 수준 지표 개발 필요	관련성
	주요 항목의 추세 및 추세의 일관성 파악이 가능하도록 시계열 자료 제공 필요	관련성 비교성
	공표 자료 포맷 다양화 필요	접근성/ 명확성
기타	원자료, DB 제공 필요	접근성/ 명확성
	공표 자료 내 부적절한 용어 수정 필요	관련성

나. 이용자 만족도 조사

이용자 만족도 조사항목은 크게 '이용 실태', '통계 만족도', '공표 범위의 세분화 필요성', '원자료 제공 필요성', '재이용 의사 여부', '품질 변화', '기타 개선 의견' 등이며 구체적인 항목은 표 3과 같다.

<표 3> 이용자 만족도 조사 항목

분류	항목
통계이용실태	1. 통계자료 이용 빈도
	2. 통계자료 이용 목적
	3. 통계자료 이용 자료 형태
	4. 통계자료 이용 경로
통계만족도	5-1. 통계자료 공표시기 적절성
	5-2. 통계자료 예고 공표 일정 준수
	5-3. 통계자료 검색의 용이성
	5-4. 통계자료 이용 시 유의사항, 관련개념, 용어정의 제공
	5-5. 통계자료의 제공정보 충분성
	5-6. 통계자료의 제공정보 다양성
	5-7. 통계자료의 신뢰성
	5-8. 마이크로데이터 이용의 편리성
	5-9. 지출비용에 대한 만족도
	5-10. 통계자료 시계열 비교 편리성
	5-11. 통계자료 국가 간 비교 편리성
	5-12. 통계자료 관련 통계 간 비교 편리성
	5-13. 급변하는 IT산업 변화에 대한 적응성
	5-14. 통계에 대한 전반적 만족 수준
공표 범위	6. 공표 범위 세분화 필요성
원자료 제공	7. 원자료 제공 필요성
재이용 의사	8. 향후 재이용 의사
품질 변화	9. 통계의 품질 변화
개선 의견	10. 개선 의견 (개방형)
응답자 특성	성별, 연령, 직업, 통계활용비중 등

응답자 특성에서 특기할 만한 것은 응답자의 근무처 및 직업 구성을 살펴보면 공공기관 이용자(85.0%) 및 학계 이용자(8.3%)가 대부분을 차지하고 있고, 통계활용비중별로 보면 보통 이상인 비율이 98.3%이었던 점이다. 즉 응답자들이 본 통계에 대해 관심을 갖고 적극적으로 활용하는 관련 분야 이용자들이기 때문에 이들이 제공한 설문 결과에 의한 이용자 만족도 조사가 충분히 설득적임을 뜻한다.

① 통계 이용 실태: 이용빈도가 주 1회 이상인 적극적인 이용자의 비율이 25%에 이르고, 분기 1회 이상 이용하는 응답자의 비율이 73.3%에 이르는 것으로 파악되었다. 주된 활용 용도는 자료의 특성 상 정책 수립 및 평가인 경우가 51.7%로 가장 많았고, 이용 형태는 간행물 및 보고서의 경우가 70.0%로 가장 많았다. 상세한 조사 결과는 <표 4>를 참조하기 바란다. 이상 조사된 이용 실태 조사 결과는 앞서 이루어진 주 이용자 층(엄밀하게는 응답자)이 본 통계를 적극적으로 활용하는 관련 분야 종사자들이라는 응답자 특성 분석 결과와 대체로 합치되는 결과이다. 한편, 현재 공식적으로 원자료 제공을 하지 않고 있기 때문에 마이크로 데이터 형태의 이용 비율이 거의 없지만, 향후 원자료 제공 서비스가 개시되면 상당히 많은 이용자가 이용하게 될 것으로 예상된다.

<표 4> 이용실태 분석

이용 빈도	주 1회 이상	25.0%
	월 1회 정도	23.3%
	분기 1회 정도	25.0%
	반기 1회 정도	5.0%
	연 1회 정도 또는 더 드물게 이용	21.7%
활용 용도	정책수립 및 평가	51.7%
	연구, 학술, 학습활동	18.3%
	사업 및 경영계획 수립	30.0%
이용 형태	간행물, 보고서	70.0%
	DB 자료	21.7%
	마이크로데이터	1.7%
	보도자료	6.7%

② 통계 만족도 조사: 각 조사항목별 만족도(5-1~5-13)가 전반적 만족도(5-14)에 미치는 영향을 고려한 항목별 중요도(가중치)를 산정하여(<그림 1> 참조) 가중합 형태로 통계에 대한 종합만족도를 산출하였다(<표 5> 참조). 구체적인 산출 공식은 아래와 같다.

○ 항목중요도 산정: 항목중요도 $w_i = \frac{r_i^2}{\sum_j r_j^2}$.

단, r_i 는 i 번째 항목(5- i)과 전반적 만족도(5-14) 간의 상관계수이다.

○ 종합만족도 계산: 종합만족도 $S = \sum_i w_i S_i$.

단, S_i 는 i 번째 항목(5- i)의 만족도이다.



<그림 1> 항목별 중요도 비교

눈에 띄는 점은 '5-4. 통계자료 이용 시 유의사항, 관련개념, 용어정의 제공', '5-9. 지출비용에 대한 만족도' 등의 항목에 대해 느끼는 항목중요도가 각각 0.073과 0.059로 항목중요도 평균 0.091을 한참 하회하는 낮은 수치를 보였으나, 이들 항목의 항목별만족도 수준은 각각 4.250과 4.617로 항목별만족도 평균인 4.239를 상회하였다는 점이다. 이는 응답자 특성 분석에서 살펴본 바와 같이 본 통계의 응답자들이 해당 통계 관련 분야의 고유 지식을 가진 전문가들로서 이들 항목에 대해 별 문제를 느끼거나 자체 해결 가능한 층이기 때문인 것으로 풀이된다. 응답자 중 상당수가 내부 이용자라는 점도 중요한 원인일 것이다.

<표 5> 종합만족도 산출 결과

항목	상관 계수 r_i	항목 중요도 w_i	만족도 S_i	가중 점수 $w_i S_i$
5-1. 통계자료 공표시기 적절성	0.672	0.102	4.333	0.441
5-2. 통계자료 예고 공표 일정 준수	0.598	0.080	4.283	0.345
5-3. 통계자료 검색의 용이성	0.654	0.096	4.233	0.407
5-4. 통계자료 이용 시 유의사항/관련개념, 용어정의 제공	0.570	0.073	4.250	0.311
5-5. 통계자료의 제공정보 충분성	0.666	0.100	4.250	0.424
5-6. 통계자료의 제공정보 다양성	0.654	0.096	4.233	0.407
5-7. 통계자료의 신뢰성	0.627	0.088	4.283	0.378
5-8. 마이크로데이터 이용의 편리성	0.617	0.086	3.851	0.329
5-9. 지출비용에 대한 만족도	0.512	0.059	4.617	0.272
5-10. 통계자료 시계열 비교 편리성	0.679	0.104	4.200	0.436
5-11. 통계자료 국가 간 비교 편리성				
5-12. 통계자료 관련 통계 간 비교 편리성				
5-13. 급변하는 IT산업 변화에 대한 적응성	0.718	0.116	4.100	0.476
평균		0.091	4.239	
합계		1.000		4.226 (종합 만족도)

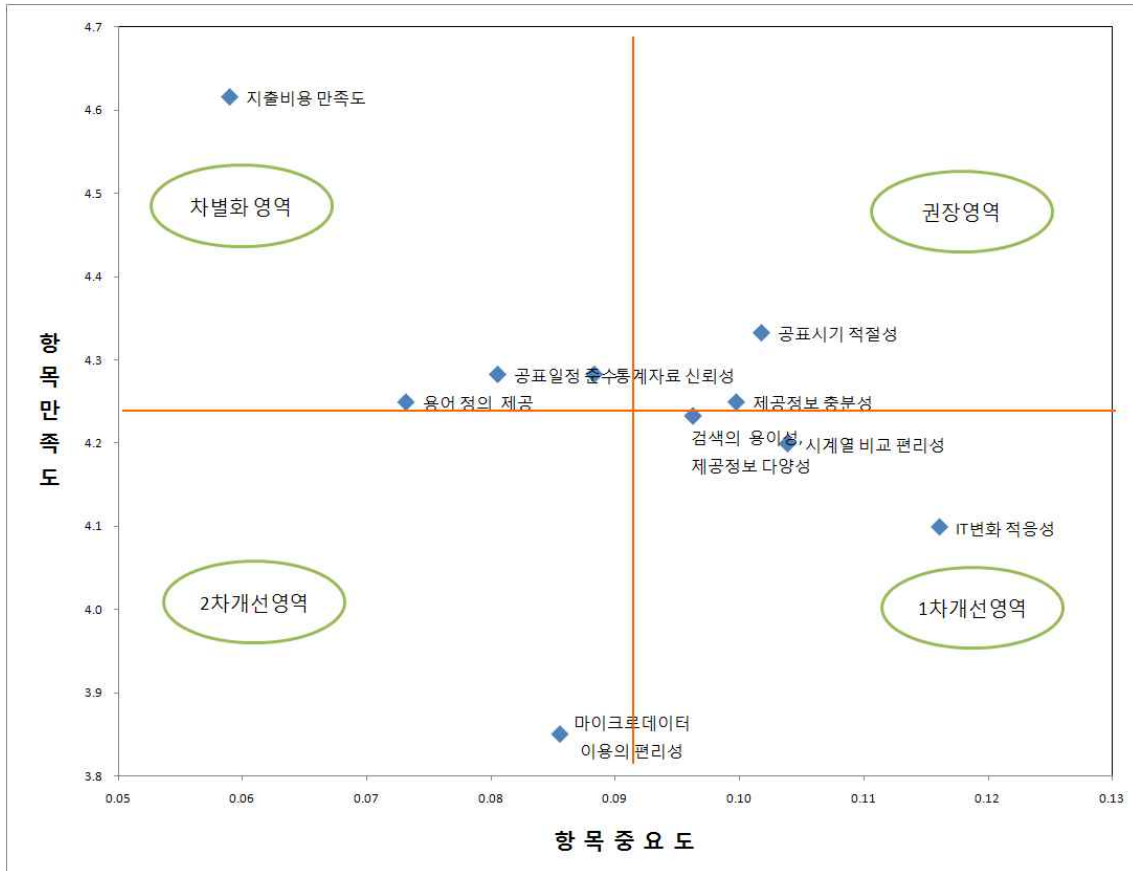
<표 6> 종합만족도 및 전반적(체감) 만족도 비교표

응답자 수	종합만족도	전반적(체감) 만족도
60명	4.226점	4.167점

종합만족도를 보면 4.226점으로 전반적(체감) 만족도 4.167점과 거의 차이가 나지 않으며 이 정도면 두 가지 만족도 모두 상당히 높은 수준인 것으로 판단된다(<표 6> 참조). 그러나 이렇게 높은 만족도 수치들이 이용자 만족도 수준이 실질적으로 높음을 의미하지는 않는 것으로 판단할 수 있다. 앞에서 이미 언급된 바대로 응답자들이 대체로 관련 분야 종사자 혹은 내부 이용자이기 때문에 이용 과정에서 발생한 요구 사항 및 불만 사항들을 스스로 혹은 우회하여 해결하는 경우가 많기 때문에 문제에 대해 체감하는 불편한 정도가 반감되기 마련이다. 또한 품질 개선의 여지가 충분함에도 비교 가능한 통계가 국내외에 전무하기 때문에 그것을 인지하지 못한 영향도 큰 것으로 보인다. 다음의 개선 과제 도출 과정 및 결과를 살펴보면 보다 또렷이 이러한 문제가 드러난다.

항목별 만족도와 중요도를 순서쌍으로 묶어 가로축은 중요도, 세로축이 만족도인 좌표평면 상에 표시한 것을 포트폴리오 행렬(portfolio matrix)이라고 한다. 각 항목의 순서쌍이 만족도 평균과 중요도 평균을 중심으로 나뉘는 4개의 영역 중 어디에 속하느냐에 따라 긴급히 개선해야 할 항목(1차 개선영역), 점차 개선해야 할 항목(2차 개선영역), 차별화(차별화 영역)하거나 권장해야 할 항목(권장 영역)으로 분류할 수 있다. 본 통계의 경우 분석에 포함시킨 11개 측정항목 중 1차 개선영역에 속한 항목은 '5-3. 통계 자료 검색의 용이성', '5-6. 통계 자료의 제공 정보 다양성', '5-10. 통계자료 시계열 비교 편리성', '5-13. 급변하는 IT산업 변화에 대한 적응성' 등 4개 항목이었고, '5-8. 마이크로데이터 이용의 편리성'이 2차 개선영역으로 분류되었다. 1차 개선영역에 속한 항목 중 '5-3. 통계 자료 검색의 용이성'과 '5-6. 통계 자료의 제공 정보 다양성'의 항목별 만족도 값은 4.233으로 만족도의 평균값인 4.239와 거의 차이가 나지 않아 상대적으로 개선의 절박성이 덜한 것으로 볼 수 있지만, '5-10. 통계자료 시계열 비교 편리성'과 '5-13. 급변하는 IT산업 변화에 대한 적응성'은 중요도 및 만족도 평균값으로 이루어진 기준점 (4.239, 0.091)으로부터 상당한 거리를 가지는 1차 개선과제로 도출되었다. 이는 본 통계로부터 시공간적으로 동적(dynamic)이고 적응적(adaptive)인 정보를 제공받고

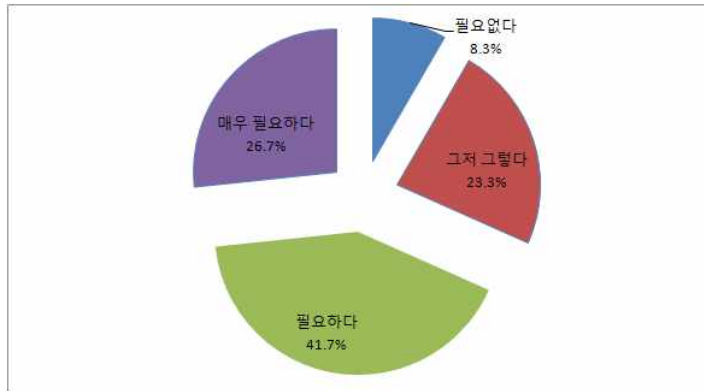
싶어 하는 이용자들의 요구가 표출된 것으로 생각된다. 2차 개선과제로 도출된 '5-8. 마이크로데이터 이용의 편리성' 항목의 만족도 3.851는 만족도 평균 값인 4.239에 한참 미치지 못하는 수준의 값으로, 비록 2차 개선영역에 속해 있지만 이용자의 요구가 상당한 항목으로 판단된다.



<그림 2> 포트폴리오 행렬

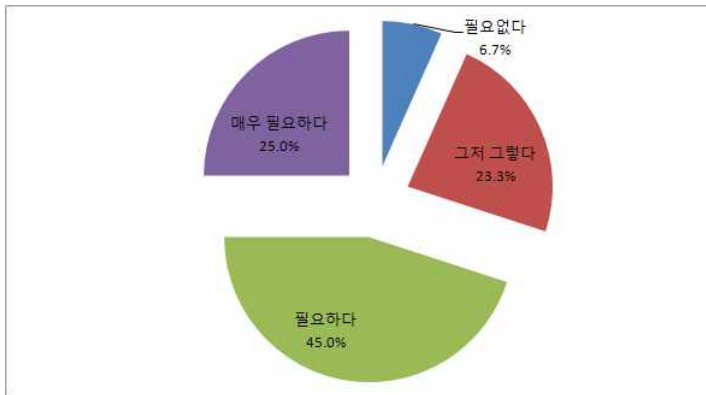
③ 공표범위 세분화: 공표 범위를 세분화할 필요성에 대해 68.3%가 '필요하다(41.7%)' 혹은 '매우 필요하다(26.7%)'라고 응답해 이용자들이 공표범위의 세분화 요구가 상당한 것으로 나타났다. 이용자들이 원하는 세분화의 수준과 세분화해 제공해 주기 바라는 자료 항목을 파악해 조사 과정 및 공표에 반영할 필요가 있다.

④ 원자료 제공: 원자료 제공의 필요성에 대해서는 70.0%가 '필요하다' 혹은 '매우 필요하다'라고 응답했다. 본 통계의 특성 상당한 어려움이 있을 것으로 예상되지만 이용자들의 수요가 있는 한 가능한



<그림 3> 공표범위 세분화 필요성

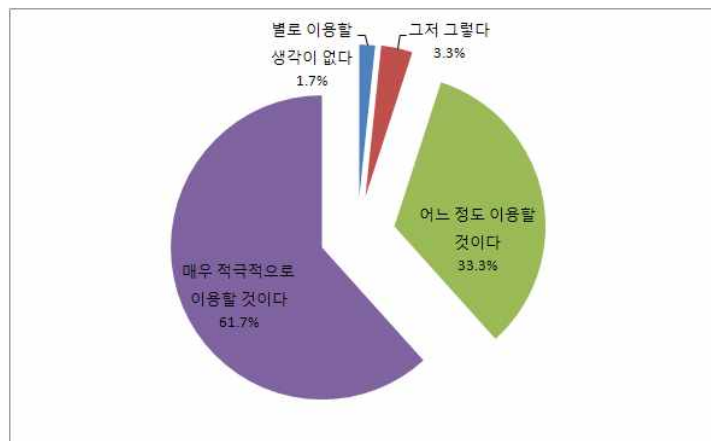
범위 내에서 원자료를 제공할 수 있는 방안을 강구할 필요가 있다.



<그림 4> 원자료 제공 필요성

⑤ 재이용 의사: 본 통계를 매우 적극적으로 재이용하겠다고 응답한 비율이 61.7%, 어느 정도 이용하겠다는 응답의 비율이 33.3%로 재이용 의사가

높은 통계이다. 이용자의 만족도가 그만큼 높다고 볼 수도 있지만, 이 분야에서 국내에서 생산되는 통계 중 거의 유일한 승인통계임을 고려할 때 지속적으로 관심을 갖고 발전시켜야 할 필요를 느끼게 해주는 조사 결과이다.



<그림 5> 재이용 의사

⑥ 품질 변화: 본 통계의 품질이 향상되었거나 매우 향상되었다고 응답한 비율이 91.7%인 반면 부정적인

응답은 전혀 없었다. 이용자들이 본 통계 작성 과정의 어려움에 대해 잘 이해하고 있는 관련 분야 종사자 그룹임을 참작하더라도 이용자들이 품질이 개선되어가고 있음을 실질적으로 체감하고 있으며, 향후 품질 향상이 지속적으로 이루어질 것에 대한 기대가 함축되어 있는 것이라 볼 수 있다.

3. 세부 작성절차별 체계

3.1 진단 개요

통계 작성과정을 '1. 통계작성기획', '2. 통계설계', '3.자료수집', '4. 자료입력 및처리', '5. 자료분석및품질평가', '6. 문서화및자료제공', '7. 사후관리' 등으로 구분하여 각 절차별 품질을 진단했다. 세부작성절차별 점검표를 이용해 내부진단, 외부진단 그리고 연구진진단으로 나누어 진단이 진행되었다. 내부진단은 통계생산기관 담당자가 스스로 통계품질진단서를 작성하면서 진단하는 것이다. 품질지표 및 지표별 평가요소들을 면밀히 검토하고 평가요소의 이행상태를 점검한 후 통계품질 상태를 5점 척도로 평가하여 품질점검표를 작성한다. 외부진단은 진단 대상 통계 또는 품질관리 분야의 전문가 2인이 실시하는 진단을 뜻한다. 제출된 내부진단 결과와 근거 자료, 그리고 본인이 갖고 있는 대상 통계에 대한 지식을 바탕으로 내부진단 방법과 동일하게 품질점검표를 작성한다. 마지막으로 내부진단 결과 및 근거 자료, 2인의 외부진단 결과, 그리고 통계품질진단 과정에서 습득한 정보 및 지식을 바탕으로 연구진진단이 이루어지며, 점검표 작성 방법은 내/외부 진단과 동일하다.

본 통계의 세부작성절차별 진단을 위해 실시한 내부진단, 외부진단, 연구진진단 결과를 정리하면 표 7과 같다. 점수 대신 회색으로 채워진 셀은 진단자가 '해당없음'으로 처리했음을 뜻한다. 각 평가 위원 간에 점수 및 의견에 많은 차이를 보인 진단 항목들이 있는데 정리하면 다음과 같다.

'1. 통계작성기획'의 1-4항에서 연구진이 해당없음으로 처리한 것은 본 통계

에 대한 국내외 표준화된 기준이 없기 때문이다. 반면에 내·외부 진단은 연구진과는 달리 이 진단 항목을 광의로 해석하여 본 통계 작성 시 승인통계들이 일반적으로 사용하고 있는 정의, 기준 및 분류체계를 준수하고 있는 점을 보고 비교적 우수한 점수를 부여했다.

‘2. 통계설계’의 2-1항에서 내·외부 진단 점수와 진단팀 점수에 많은 차이를 보이고 있는데, 조사모집단과 목표모집단이 일치하지 않는 문제에 대한 인식 차이에서 기인한 것으로 판단된다. ‘2-5. 표본규모의 적정성’ 역시 비슷한 점수 분포를 보이는데, 내/외부 진단의 경우 본 통계 작성을 위한 예산 수준을 고려해 비교적 후한 점수를 받았던 것으로 확인되었으나 연구진 입장에서는 표본 규모가 절대적으로 너무 작다는 점을 간과할 수 없는 중요한 문제로 생각하여 낮은 점수를 부여하였다.

‘5. 자료분석 및 품질평가’의 경우 ‘해당없음’으로 처리된 경우가 많다. 5-4항과 5-8항의 경우 특히 그러한데 본 통계가 중요한 통계작성방법 등의 변경을 경험한 적이 없었고 잠정치를 공표하지 않기 때문이며, 외부진단 2는 같은 이유 때문에 이들 항목에 중립적인 점수인 3점을 부여한 것이다.

6-5항은 조사 항목 중 공표 시 누락되는 것이 있으나 본 통계의 특수한 사정을 고려해 내·외부 진단에서 ‘해당없음’ 혹은 높은 점수를 받았으나, 사안의 중요성을 감안한 연구진 진단에서 낮은 점수를 받았다. 6-8항은 공표 절차가 부분적으로 미비된 사항이 있어 내부진단에서 해당사항으로 처리했으나 외부 및 연구진 진단에서 비교적 낮은 점수를 받았다.

<표 7> 세부작성절차별 점검 지표 및 점수

작성 절차별 분류	작성절차별 품질진단 지표	내부진단	외부진단 1	외부진단 2	연구진진단
1. 계 작성 기획	1-1. 통계작성 목적 제시	5	5	4	5
	1-2. 이용자의 요구 및 이용실태를 파악	5	5	3	4
	1-3. 개념, 용어, 분류체계 등의 타당성 검토	5	5	4	5
	1-4. 표준화된 정의, 기준 및 분류체계 준수	5	4	4	
	1-5. 통계작성 개편작업의 적절성	5	5	4	4
2. 통 계 설계	2-1. 통계작성대상의 명확한 정의	5	5	5	2
	2-2. 응답자가 이해하기 쉽고 작성하기 편리한 조사표 설계	5	3	5	4
	2-3. 조사항목을 추가, 변경 시 사전검토	5	5	2	3
	2-4. 표본오차 관련 지표 작성	5	5	4	4
	2-5. 표본규모의 적정성	4	5	4	2
	2-6. 표본수준 유지를 위한 적절한 표본관리	5	5	4	5
	2-7. 표본설계에 대한 상세 정보를 제공	5	5	5	4
3. 자 료 수 집	3-1. 조사직원을 위한 지침 제공	5	5	4	5
	3-2. 조사직원에 대한 체계적인 교육 실시	5	5	4	5
	3-3. 현장조사 실시에 대한 체계적인 관리	4	4	4	4
	3-4. 조사기획자가 현장조사에 대한 단계별 업무량 파악 여부	5	5	3	5
	3-5. 조사기획자가 조사직원의 조사관련 지식 숙지 여부 파악	4	4	3	4
	3-6. 현장조사 질의사항의 시의적절한 처리 및 직원 간 공유	5	3	3	5
4. 자 료 입 력 및 처 리	4-1. 자료 입력을 위한 표준화된 체계	5	5	5	5
	4-2. 체계적 자료 내용검토(에디팅)작업	5	5	4	5
	4-3. 무응답 실태를 파악 및 분석	5	5	5	4
	4-4. 현장조사, 집계, 분석 단계까지 내용검토 절차	5	5	4	5
5. 자 료 분 석 및 품 평 가	5-1. 관련통계 등과의 비교분석을 통해 결과를 검증	5	5	4	
	5-2. 시계열자료의 연속성	5	5	3	
	5-3. 집계결과와 통계표간 일관성	5	5	3	3
	5-4. 경제·사회현상, 통계작성법 변경 등이 미치는 영향 분석			3	
	5-5. 모수 추정절차의 적정성	5	5	4	5
	5-6. 중간 통계자료에 대한 검증	4	5	3	4
	5-7. 최종 통계자료에 대한 검증	5	5	4	5
	5-8. 잠정치, 확정치 간의 불일치 분석 및 관리			3	
6. 문 서 화 자 료 제 공	6-1. 통계작성과 관련 각종 자료 문서화	5	5	4	5
	6-2. 간행물 수록자료에 대한 오류 점검	4	4	3	4
	6-3. 간행물에 설명자료를 수록	5	5	4	4
	6-4. 개편 작업 후 개편내용 공개	5	5	3	5
	6-5. 조사한 항목을 모두 공표			5	2
	6-6. 공표시 모든 이용자가 조사결과를 동시에 이용	4	4	5	4
	6-7. 결과 자료를 적절한 시점에 공표	4	4	4	3
	6-8. 결과 자료의 공표절차 준수		2	3	2
	6-9. 다양한 매체를 이용한 결과자료 제공	4	5	4	3
	6-10. 자료제공 시 개인 비밀보호 장치	5	5	4	5
	6-11. 동일 주제의 다른 통계자료와 비교 분석	4	4	3	
7. 사 후 관 리	7-1. 새로운 정보요구에 신속 대응 가능한 통계작성 체계	5	5	3	5
	7-2. 고품질 통계 생산을 위한 전문성 유지 및 개선 노력	5	5	3	5
	7-3. 통계 작성방법에 대하여 타당성 검토	4	4	3	3
	7-4. 합리적이고 효율적으로 통계를 작성하기 위한 품질관리	5	5	3	4

3.2 현황 및 문제점

가. 작성절차별 통계 품질

각 항목의 점수를 절차별분류에 따라 평균값을 구하여 분석을 실시하였다 (표 8, 그림 6 참조). '1. 통계작성기획', '4. 자료입력및처리' 등의 품질은 대체로 양호하다는 평가를 받았으나, '2. 통계설계', '3. 자료수집', '5. 자료분석및품질평가'와 '6. 문서화및자료제공', '7. 사후관리' 등의 부문은 상대적으로 낮은 점수를 받았다.

'2. 통계설계' 부문이 다소 낮은 수준의 평가를 받은 것은 조사표가 복잡하고 조사항목이 많으며 표본규모가 작기 때문이었다. 조사표가 복잡하고 조사항목이 많아진 것은 본 통계 관련 조사 항목 외에 작성기관이 작성하는 다른 통계의 조사 항목까지 함께 조사표에 포함시키고 있기 때문인 것으로 파악되었다. 그러나 '민간기업정보보호실태조사'가 지닌 승인통계의 지위에 따라 여러 혜택을 받고 있는 점을 감안하면 조사표 설계 시 본 통계 관련 항목의 조사가 피해를 보지 않도록 배려할 필요가 있다. 표본규모 문제는 이미 이용자만족도 분석 과정에서 제기된 바 있는 문제이므로 여기에서는 논의하지 않기로 한다.

'3. 자료수집' 절차는 작성기관에서 조사 현장에 대한 관리를 적극적으로 하지 않고 있는 것 때문에 다소 낮은 평가를 받았다. 이 부분은 자료수집정확성 부문 진단에서 관련 내용을 다시 자세히 다루므로 연구보고서 내에서 내용의 중복을 피하기 위해 여기서는 더 이상 논의하지 않는다.

'5. 자료분석및품질평가'의 경우 중간통계자료의 검증 절차 관련 항목에서 좋지 않은 점수를 받아 비교적 낮은 평가를 받았다. 가집계 자료에 대한 점검이 이루어지고 있는 것으로 파악되었으나, 그 절차의 체계가 아직 문서화되어 있지 않아 담당자가 바뀌는 등 환경의 변화가 발생하는 경우 점검의 실

효성을 장담할 수 없다. 중간 점검 절차를 위한 매뉴얼을 작성할 필요가 있다.

‘6. 문서화및자료제공’에 대한 평가가 가장 나빴는데, 간행물 오류 점검이 소홀한 점, 조사 항목 중 공표 시 누락시키는 항목이 많은 점, 공표 절차가 제대로 지켜지지 않은 점 등이 주된 원인이다. 간행물 오류 점검과 공표 절차 준수의 문제는 작성기관의 의지만 있으면 즉시 해결 가능한 것이므로, 차년도 보고서부터는 이들 문제의 해결을 위한 절차가 바로 진행되어야 한다. 그러나 조사 항목 중 공표 시 누락시키는 항목이 많은 점을 해결하는 것은 간단한 문제가 아니다. 즉 작은 표본 규모에 따른 오차 관리의 어려움, 정보보안 관련 항목의 민감성, 무응답 발생 등의 문제 등을 면밀히 따져야 한다. 어쨌든 조사 비용 대비 공표 자료의 효용에 대한 효율성을 따져볼 때 공표 항목의 범위를 확대해야 한다는 원칙적 논점은 분명해 보인다.

‘7. 사후관리’의 점수도 비교적 낮은 편인데, 통계자료의 활용에 대한 모니터링이 거의 이루어지지 않고 있고 통계 생산 절차의 효율성에 대한 체계적 검토가 이루어지지 않고 있는 것이 주된 원인이었다. 자료의 활용도를 관리해야 한다는 것에 대한 인식이 그간 부족했고, 자료 생산 자체에 급급하여 생산 절차상의 효율성 문제에까지 관심의 범위를 넓히지 못했던 것으로 보인다.

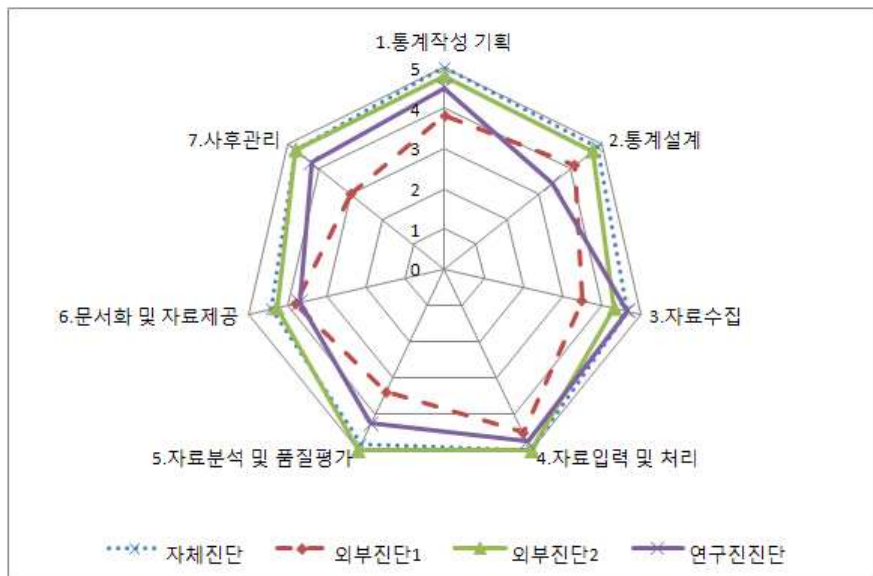
나. 품질차원별 통계 품질

각 항목을 품질차원별로 분류한 후 품질차원별로 평균값을 구해 분석을 실시하였다. 세부작성절차별 점검을 통해 본 통계의 품질차원별 품질 중 가장 좋지 않은 것은 시의성으로 드러났다. 이를 개선하려면 공표 절차 특히 공표 일정을 공시하고 공시된 일정을 준수하는 것이 필요하다. 일관성 차원의 품질이 다소 낮은 점수를 받았는데 비교 가능한 다른 통계가 거의 없기 때문에 깊이 있는 논의가 불가능하다. 나머지 차원의 품질은 비교적 양호한 것으로

로 나타났다.

<표 8> 작성절차별 품질진단

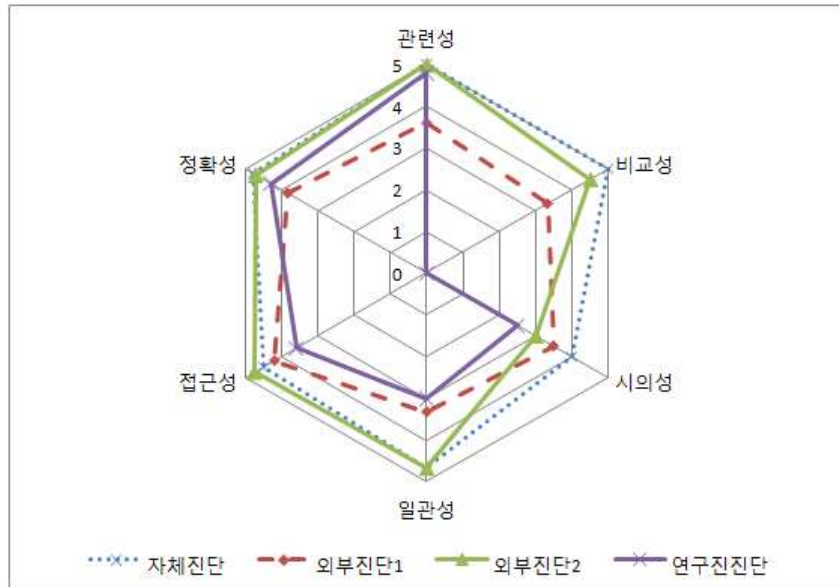
	1. 통계 작성 기획	2. 통계 설계	3. 자료 수집	4. 자료 입력 및 처리	5. 자료 분석 및 품질 평가	6. 문서화 및 자료 제공	7. 사후 관리
내부진단	5.00	4.86	4.67	5.00	4.83	4.44	4.75
외부진단1	3.80	4.14	3.50	4.50	3.38	3.82	3.00
외부진단2	4.80	4.71	4.33	5.00	5.00	4.30	4.75
연구진진단	4.50	3.43	4.67	4.75	4.25	3.70	4.25
평균	4.53	4.29	4.29	4.81	4.37	4.07	4.19



<그림 6> 작성절차별 품질진단

<표 9> 품질차원별 품질진단

	관련성	비교성	시의성	일관성	접근성	정확성
내부진단	5.00	5.00	4.00	4.67	4.50	4.79
외부진단1	3.60	3.33	3.50	3.33	4.20	3.84
외부진단2	5.00	4.50	3.00	4.67	4.75	4.71
연구진진단	4.80		2.50	3.00	3.60	4.29
평균	4.60	4.28	3.25	3.92	4.26	4.41



<그림 7> 품질차원별 품질진단

다. 표본 설계 진단

<표 10> 진단대상 통계 개요

조사명	민간기업정보보호실태조사		조사주관	방송통신위원회
표본설계 방식	자체설계 ()		외부용역 (○) 【기관명 : 한국인터넷진흥원】	
조사목적	국내 기업 및 기관의 정보보호 실태 및 문제점에 대한 현황 파악과 원인 분석을 통하여 정보보호 인식 제고 및 수준 향상을 위한 정책 개발			
조사대상	전국 종사자수 5인이상, 네트워크로 연결된 PC를 1대 이상 보유한 사업체 2,234개			
조사항목	정보보호환경, 정보보호대책, 개인정보보호조치, 침해 사고 대응 및 스팸통제, 침해사고피해 등 정보보호 실태에 대한 조사			
조사방법	방문면접조사	조사시기 (조사주기)	2009. 11. 10 ~ 2009. 12. 30 (매년)	

(1) 표본설계 현황

- 모집단 : 종사자수가 5인 이상이고 네트워크가 구축된 컴퓨터가 1대 이상인 사업체
- 표본추출틀
 1. 사업체기초통계조사에서 대상 사업체
 2. 2009년 정보화통계조사 사업체 중 네트워크 구축 사업체 : 구축률을 활용 => 한국표준산업분류의 의한 20개의 대분류 중 가사서비스업, 국제 및 외국기관, 자동차관련도·소매업을 제외한 18개 산업에 속하

고 종사자 수가 5명 이상인 사업체 531,345개 중 네트워크에 연결된 컴퓨터를 1대 이상 보유한 사업체 308,529개 업체

- 목표오차 : 95% 신뢰수준에서 총계의 허용오차를 $\pm 2.1\%$ 이내로 함
- 표본규모 : 2,234개 사업체
- 표본배정 및 추출방식
 - 2008년 사업체기초통계조사 자료에서 5인 이상 사업체들을 업종별(10종) \times 규모별(5종)로 총 50개의 층으로 분류
 - 2009년 정보화통계조사 자료를 이용하여 모집단의 분포를 파악
 - 예산에 의해 정해진 약 2,500개의 전체 표본규모를 업종별 규모에 따라 비례 배분
 - 업종별과 규모별로 2단 층화한 후 각 사업체들을 지역별로 정렬하여 계통추출을 기본으로 함
 - 절사추출 : 종사자 수가 1,000명 이상인 사업체는 전수조사 실시, 종사자 수 250~99명인 사업체는 일부 전수조사.
 - 과표본추출 : 50개 층별로 최소 20~30개 우선 할당.
 - 비례할당 : 업종 및 규모별 조사 가능성을 고려하여 배분.
- 가중치
 - 본 조사는 층별 비례할당을 하지 않았기에 모집단과 표본간의 편차 최소화를 위해 가중치를 이용한 사후층화를 이용함.
 - 비율이 낮은 250이상 규모의 사업체는 통합 계산.
 - 가중치 계산
$$W_{gh} = \frac{N_{gh}}{n_{gh}} \cdot \frac{1}{R_{gh}}$$
 - N_{gh} = g 업종의 h 규모의 조사모집단의 크기
 - n_{gh} = g 업종의 h 규모의 층에 할당된 표본크기
 - R_{gh} = g 업종의 h 규모의 층의 응답률
- 추정식
 - 허용오차
$$e = Z_{\alpha/2} \times \sqrt{P(1-P)/n(1-n/N)}$$
 - N = 모집단 크기

n = 표본 크기

· 모수 추정 $\hat{Y} = \sum_{h=1}^L \hat{Y}_{hc} + \sum_{h=1}^L \frac{N_{hs}}{n_{hs}} \sum_{k=1}^{n_{hs}} y_{hsk}$

L = 층의 개수 (업종×규모)

n_{hs} = h 층의 표본 크기

y_{hsk} = h 층 표본의 k 번째 관측값

\hat{Y}_{hc} = 전수조사에서 각 층의 총계에 대한 추정량

- 표본오차: SAS SURVEYMEANS에서 제공
- 공표 범위 : 전국단위 업종별, 규모별
- 주요 공표 항목 : 정보보호정책수립비율, 전담조직 운영현황비율, 인지도 (비율), 정보보호교육실시 비율, 교육시간대 비율, 투자비율, 서비스 도입 비율, 침해사고비율 등

(2) 문제점

① 표본규모 문제: 표본규모는 주요 항목에 대한 표본오차를 기준으로 정해 지는데 본 통계는 '정보화 투자 대비 정보보호 투자 비율'항목을 기준으로 전국 단위의 표본오차를 구하였다. '지출없음'과 '1%미만'의 셀은 표본오차 (RSE)가 5% 미만으로 나타났으나, 전체적으로는 5%를 훨씬 초과하고 있어 표본규모를 증가시키는 것이 바람직하다고 판단된다(<표 11> 참조). '정보화 투자 대비 정보보호 투자 비율'외의 다양한 항목을 동시에 고려해 표본오차를 검토하는 것이 필요하다.

<표 11> '정보화 투자 대비 정보보호 투자 비율' 항목의 표본오차(RSE)

	지출 없음	1% 미만	1%~3% 미만	3%~5% 미만	5%~7% 미만	7%~10% 미만	10% 이상
RSE	1.71	4.91	6.07	8.21	14.59	12.03	16.77

② 부적절한 용어 사용: 본 조사의 제목은 '민간기업정보보호실태조사'이다. 따라서 민간기업을 기준으로 조사 결과를 보여주는 것이 바람직하다. 그러나 모집단과 표본들을 살펴 보면 표본추출단위가 사업체 단위로 되어 있어 모든 조사의 집계 결과는 사업체 단위로 이루어진다. 따라서 '민간기업..'이라기 보다는 '민간사업체..'로 보는 것이 바람직하다. 표본설계보고서에서도 정확한 용어의 사용이 필요하다. 예를 들면 허용오차와 표본오차의 용어가 혼용 내지는 오용되고 있다. 그리고 통계결과표는 총계 추정치를 제시하고 있는데 이 때 총계 추정치를 사례수로 표기하는 것은 적절치 않다.

③ 표본설계보고서 문제: 표본설계는 외부용역에 의해 이루어지고 있으며 표본설계내역을 별도의 보고서 형태로 받지 않고 대부분 표본리스트와 간단한 설계내역만을 받고 있다. 이렇게 되면 담당자들의 표본설계 및 추정에 관한 내용에 대한 이해도는 피상적일 수밖에 없으며, 추정 과정 및 표본오차의 검토 단계가 소홀해지게 된다. 표본설계내역을 보고서 형식으로 보관하고, 추정 및 마지막 표본오차 단계까지 지속적인 자문이 이루어지는 것이 바람직하다.

④ 통계표 일치성 문제: 통계표의 기준이 같을 경우에는 통계표들의 전체 합계의 값은 같아야 하는 것이 일반적이거나, 본 조사의 보고서는 전체 값이 일정하지 않다. 이는 무응답 및 결측값 처리를 전혀 고려하지 않았기 때문으로 보인다.

4. 자료수집의 정확성

4.1 진단 개요

자료수집정확성 진단을 위해 현장 방문을 실시하였다. 조사기획자 1인, 조사원관리자 3인, 조사원 3인과 응답자 3인을 대상으로 면담을 실시했다.

4.2 현황 및 문제점

가. 조사기획자 점검

(1) 자료수집 방법의 적절성

현재 사용하고 있는 자료수집방법은 대면조사를 원칙으로 하고 있다. 부득이하게 응답자가 자기기입식 조사를 원하는 경우에는 예외적으로 진행하되 설문지 회수 시 현장에서 응답 내용의 누락이나 오기 등을 재점검하고 있다.

조사표 회수율은 지역이나 업종에 따라 차이가 있으나, 원표본 유지율 30 ~ 40% 수준이다. 전수조사 층과 표본조사 층을 구분하여 관리하고, 조사가 어려운 전수조사 층이나 업종의 경우에는 별도의 전문상담원을 통하여 대상자를 섭외하여 조사가 원활히 진행되도록 하고 있다.

무응답률 문제는 응답자가 현황 파악이 되지 않아 발생하는 문항 무응답이 일부 있으나 1% 내외로 결과에 영향을 주는 정도는 아니다. 항목 무응답 방지를 위해 현장 점검을 강화하고는 있다. 그러나 모수 추정 시 무응답 대체는 실시하지 않고 있다.

본 조사의 조사기간은 2개월이며, 조사 예산은 2009년의 경우 약 1억2천만원이며 이 중 조사원 인건비는 한 건당 약 25,000원이 지불된다.

(2) 현장 점검 및 관리체계

조사 기획자는 표본설계, 실사관리원 관리, 통계처리, 조사결과 분석 등 조사 업무 전반에 대한 관리 및 감독을 담당하고 있다. 실사관리원과의 현장 점검 등을 통한 즉각적인 대처, 조사결과의 시계열적 확인 등의 조사 전반에 대한 관리 및 점검이 필요하므로 업무량이 많은 편이다.

현장 조사 단계에서 부실 조사를 사전에 예방하기 위해 다음과 같은 조치가 이루어진다. 실사 지도를 위한 조사지침서(조사원 업무 수행사항)에 의해 면접원들의 사전 교육이 이루어지는데, 조사항목누락, 품목내용오류 등을 방지할 수 있게 조사표를 작성 검토하도록 설명하고 있다. 또한 '연관관계분석'을 통하여 조사(입력)오류를 찾아내고 수정하고 있다. 이전 조사의 문제점을 사전에 점검하여 지침을 마련하고, 조사 결과의 시계열적 추이를 수시로 점검하여 조사결과에의 정확성을 확인하고 있다. 또한 조사 진행률에 따라 회수된 설문지를 면접원별로 점검하여 오류를 확인하고 필요시 재조사 및 추가 지침을 마련해 제공하여 지침을 보완하고 있다. 방문시간대, 면접시간, 재접촉 등에 대한 교육에서도 역시 이전 조사의 경험을 활용하여 업종 및 규모별 대처방안을 마련하고 있다.

실사 지도에서 도출된 문제점 및 착오 내용에 대해 조사 진행 전에 집체 교육과 더불어 수시 점검을 통해 재교육 및 보완교육을 실시하고 있다. 특정 사안이 발생하는 즉시 투입된 전체 면접원을 교육함으로써 조사 기준의 통일성을 확립하고 있다.

조사내용의 정확성(사전 사후 내용 검토) 확인을 위해 면접원, 실사관리원, 담당연구원간 단계별 확인 체계를 마련하여 조사 내용을 점검하고 있다. 또한 응답자 검증 및 재조사를 위하여 응답자 재추적용 연락처는 반드시 확보해야 하는 사항이며 명함 등을 첨부하도록 하고 있다.

현장 에디팅을 강화함으로써 오류를 최소화하고 있다. 현장에서 응답 내용에 애매한 경우 면접원에게 상황을 체크하도록 함으로써 사후 확인이 가능하도록 하고 있다. 현장 내용 검토 및 에디팅 지침은 이전 조사 결과를 반영한 현장 에디팅 가이드를 마련하여 제공하고 있다.

(3) 대상처/응답자(표본) 관리

표본 대체 지침은 대상처에 대한 표본대체 지침을 마련하고 있는 것과는 별도로, 1차적으로 실사 관리원에게 통보하여 지침을 받아 진행하도록 하고 있다. 최초 표본리스트와 최종 응답자리스트 비교에 따른 표본 준수 여부는 조사된 리스트 확인을 통하여 점검하고 있고, 면접원별로 원표본 준수 상황을 관리함으로써 과도한 대체가 이루어지는 면접원을 재교육하거나 교체하고 있다.

나. 조사원 관리자 점검

(1) 조사원 관리자

조사원 관리자는 조사리스트 분배, 회수 및 정리, 조사원 관리, 설문지 내용에디팅 및 내용 확인, 조사 일정 관리, 총 서울 관리자가 서울 및 기타 지방 일정 및 데이터 확인 작업, 면접원 섭외, 검증, 이관 작업, 주간 면접 진행 상황보고, 실사관리 프로그램 운영, 면접원별 설문지 배부 부수 점검, 접촉률 응답률 등 면접상황에 대한 집계, 조사대상자 선정방법, 거부, 무응답처리법, 문제발생시 조사원 처리법 등 진행과정에서 발생하는 일들의 모니터링, 보고, 기록 등 많은 일을 한다.

모든 조사원이 임시일용조사원이며 선발기준은 리스트 조사에 3년 이상의 경력을 가지고 용모단정하고 깔끔하며, 끈기가 있고 친절한 응대가 가능한 자, 정직하고 성실하며 책임감 있게 조사 수행이 가능한 자, 사교적이며 긍정적인 사고를 가진 자를 선정하며 타 회사 조사경험자는 우대해 주고 있다.

신규면접원의 실사 지도를 위해 우선 조사원으로 투입하기 전에 별도로 월 2회의 면접조사법, 진행법, 설문지작성법 등 조사를 진행하는 과정에 대해 전반적인 사항에 대한 교육을 실시하고 있다. 정규 면접원 교육은 프로젝트 형태로 담당연구원 설문지에 대해 교육(집합 교육)을 실시하고 있다. 담당연

구원 오리엔테이션이 끝나면 SV는 면접원에게 질문지 상에서 유의해야 할 사항과 조사의 가장 중요한 부분을 강조하면서 2차 오리엔테이션을 실시(컨택하는 법, 거절시 대응법, 리스트 관리법 등)하고, 진행 후 추가 문항이나 주의할 사항이 있으면 추가 교육을 실시하며 수시로 전화로 면접원에게 상황 보고를 받으면서 지침을 전달하는 형식으로 운영된다.

조사원이 조사를 진행하는 과정에서 발생한 문의 사항이나 새로운 지침이 나왔을 때 조사원 모두와 통화하여 새로운 내용에 대해서 전달하고, 이와 함께 설문지 에디팅 시 잘못 작성된 내용에 대해 조사원에게 설명해주어 똑같은 실수가 반복해서 나오지 않도록 주의시키고 있다. 그러나 조사에서 조사를 진행하면서 문항 중 모르는 부분이나 특정 문항은 본사 담당자에게 전화를 해 질문을 하거나 본사 내용으로 응답을 하라는 지침이 있지만, 현실적으로 불가능하여 설문지가 폐기되거나 무응답으로 처리되는 경우가 많았다.

조사 내용의 정확성을 위하여 조사표 배부 후 30% 정도 진행되면 1차로 수거를 해서 에디팅을 실시하고 있다. 에디팅 시 부족하거나 잘못된 내용이 있으면 면접원에게 전달해주고, 검증실에서 응답자와 통화하여 조사 진행이 정확하게 되었는지, 조사원이 바른 태도를 하였는지 등을 확인하고 있다. 또한 검증원들이 상시 근무를 하고 있어 유선으로 검증하고 있다.

조사원의 업무 수행 능력을 점검하기 위해 진행 상황을 격일로 보고받고 타 조사원과 수행부수를 비교하면서 빠르고 정확하게 진행할 수 있도록 독려하고 있다. 또한 회수된 설문지에 대해서는 검증팀에 이관을 하고 검증 시 조사원의 태도 및 설문 내용을 정확하게 진행했는지 점검하고 있다. 그 외에 컨택 시 거절율과, 설문 내용 숙지 정도, 실사 마감일 준수, 껀럽 충성도 등을 점검하기도 한다.

검증 과정에서 에러가 밝혀진 설문지에 대해서는 조사원에게 피드백해서 주의를 주고 재조사를 실시하도록 하고, 리스트 정리 및 리스트 관리, 스케줄

관리 등을 관리자가 총체적으로 파악하여 부족한 부분이 있으면 면접원을 독려하고 있다. 또한 조사원이 조사를 제대로 수행하지 못하였거나 응답자를 만났을 때 잘못된 행동을 할 경우 검증실장 주관으로 월 2회 교육 실시한다.

조사내용 착오유형에 대한 교육을 위해 조사 진행 중반에 질문지를 회수하여 미리 에디팅한 후 조사원의 착오로 잘못된 내용에 대해서 지침을 다시 내려주고 교육하고 있다. 또, 회수된 설문지를 담당 조사원관리자가 회수한 후 조사원 앞에서 에디팅을 실시하고 누락 부분이나 잘못 작성된 부분은 다시 작성하도록 하고 있다.

현장 내용 검토 지침은 조사원에게 제공하여 지침서대로 진행하도록 교육할 수 차례 진행하고 있다. 특히 설문지 특성 상 전화로 검증할 수 없는 내용이 많으므로 설문지 회수 시 그 자리에서 빠진 내용이 없는지, 질문의도에 맞게 응답이 되었는지 확인하도록 관리자 오리엔테이션 때 지시 및 교육하고 있다. 실사 중에 추가지침이 내려온 경우 재교육을 실시하고 있다.

(2) 대상처/응답자(표본) 관리

응답자의 협조율은 30퍼센트 정도로 응답 거부율이 높고, 의무 사항이 아니라는 이유로 대상자 측에서 응답을 꺼리는 경우가 많다고 한다. 이에 강제는 아니지만 국가 정책에 중요한 정보로 제공될 것이라는 점을 설명하고 선물(펜 세트)을 제공하고 있다. 선물의 수준을 높일 필요가 있다고 생각된다.

본 표본준수율은 30퍼센트 정도 된다. 리스트에 본 표본과 예비표본으로 구분하고 본 표본을 우선적으로 조사하되, 실패 시 예비표본으로 대체하도록 교육하고 있다. 최초 표본리스트와 최종응답자 리스트비교는 모두들 제공받은 리스트에 아이디어가 모두 부여되어 있어 확인 가능한 사항이라고 한다. <표 12>는 조사 전(표본설계당시)과 후의 표본의 층별 분포를 비교한 것으로, 대체로 표본설계 시의 분포가 유지된 것으로 보인다. 본 자료는 보고서

에 수록되어 있는 큰 틀의 분포만을 비교하였으나, 더욱 세분화된 모집단의 분포가 가능하여 비교하는 것이 좋겠다. 이를 위해서는 표본설계 당시의 자세한 내역을 가지고 있는 것이 바람직하다.

<표 12> 조사 전·후의 표본 층별 분포 비교

업종분류	규모분류	컴퓨터보유/ 네트워크 구축사업체		업종분류	규모분류	컴퓨터보유/ 네트워크 구축사업체	
		전	후			전	후
농림 수산업	5 ~ 9명	23	23	숙박 및 음식점업	5 ~ 9명	30	30
	10 ~ 49명	23	23		10 ~ 49명	46	46
	50 ~249명	18	18		50 ~249명	23	23
	250명 이상	2	2		250명 이상	9	9
	1000명 이상	2	2		1000명 이상	1	1
제조업	5 ~ 9명	133	133	운수 및 통신업	5 ~ 9명	56	56
	10 ~ 49명	176	176		10 ~ 49명	49	49
	50 ~249명	61	61		50 ~249명	35	35
	250명 이상	6	6		250명 이상	13	13
	1000명 이상	15	15		1000명 이상	2	2
건설업	5 ~ 9명	69	69	금융 및 보험업	5 ~ 9명	63	63
	10 ~ 49명	86	86		10 ~ 49명	105	105
	50 ~249명	30	30		50 ~249명	29	29
	250명 이상	16	16		250명 이상	5	5
	1000명 이상	6	6		1000명 이상	1	1
도매업	5 ~ 9명	80	80	부동산 및 임대업, 사업 서비스업	5 ~ 9명	117	117
	10 ~ 49명	65	65		10 ~ 49명	122	122
	50 ~249명	28	28		50 ~249명	35	35
	250명 이상	9	9		250명 이상	8	8
	1000명 이상	1	1		1000명 이상	14	14
소매업	5 ~ 9명	62	62	기타 서비스업	5 ~ 9명	209	209
	10 ~ 49명	42	42		10 ~ 49명	179	179
	50 ~249명	25	25		50 ~249명	72	72
	250명 이상	16	16		250명 이상	6	6
	1000명 이상	0	0		1000명 이상	11	11

다. 조사원 점검

(1) 업무량 및 교육

2009 '민간기업정보호실태조사'의 현장조사기간은 11월 17일부터 12월 14일 까지 한 달 가량이었으며, 실사기간 동안 1인당 업무량(조사건수)은 20건에서

30건 정도였다. 설명이 필요한 어려운 조사항목이 많고, 응답을 미루는 업체들이 많아 전화 통화와 방문의 횟수가 많아져 업무량이 많은데 비해 임금이 낮게 책정되어 있다고 한다.

조사에 투입되기 전에 조사방법 및 지침에 대해 연구원과 실사 관리원들로부터 2시간 정도의 교육을 받는다고 한다. IT전문 용어들이 많고 새로운 내용들이 많아 면접관과 응답자 모두 어려운 내용도 있었으나 교육을 통해 전반적으로 파악이 가능했다고 한다. 응답자들의 문의가 많은 만큼 자신있게 대답할 수 있도록 더 자세한 교육이 이루어지길 희망하고 있다.

조사지침서를 받아 보았고, 모르는 용어가 있는 경우가 발생하더라도 관리자에게 문의하면 대체로 해결된다고 한다. 교육 시 사용되는 교재(조사지침서)에는 조사 시 알아야 할 사항이 수록되어 있어 상당히 도움이 되지만, 여전히 조사 시 발생하는 상황에 따라 임기응변으로 해결해야 할 경우가 많다고 한다.

현장에서 작성한 (또는 회수한) 조사표의 내용을 검토해 미비한 사항을 응답자에게 확인하는 과정은 조사 현장에서 직접 하고, 회수 후 반드시 이루어지고 있다.

(2) 대상처/응답자(표본)관리

조사대상에게는 우선 e-mail을 전송하고 전화로 조사하는 내용과 이유를 설명하여 조사하며, 공문들을 활용하기도 한다. 본 조사는 조사의 내용 및 양으로 볼 때 응답이 매우 어려운 조사로 응답자들에게 간단한 선물(펜 세트)을 제공하고 있다. 표본리스트(대체표본포함)를 50% 이상 준수하고 관련 사항은 관리자와 상의하고 있다.

라. 응답자 점검

조사에 참여한 3명을 면담하였다. 조사 항목에 사용된 문제에 대해서는 응답

자들이 이해되는 부분도 있고 이해되지 않는 부분도 있다고 하였으나, 조사원들의 설명으로 인해 대부분 이해가 가능했다고 한다. 설문응답 시간은 10분에서 길게는 40분까지 걸렸다고 한다.

‘선생님이 응답한 조사의 통계결과가 언론 등에서 발표한 것을 보거나 들으신 적이 있습니까’의 질문에는 언론 등에서 발표한 것은 보지 못하였지만, 추후 응답을 요청할 시 응답에 응할 의사가 있다고 하였다.

5. 통계자료 서비스 진단

5.1 진단 개요

통계자료 서비스 진단은 평가대상 통계의 보도자료 및 보고서에 대한 점검을 통하여 주로 발생하는 오류의 유형 및 빈도에 대한 자료를 작성하고 재발 방지 방안을 모색하고, 이용자들이 간행물을 이용할 때 필요한 기본정보가 충실히 제공되고 있는지를 점검하고 보완하기 위한 것이다. 통계청에서 제공한 공표자료 오류 점검표와 이용자 편의사항 점검표를 이용하여 2009 보고서에 대해 진단을 실시하였다.

5.2 현황 및 문제점

가. 공표자료 오류 점검

공표자료 오류 점검은 수치자료, 통계표 형식 및 내용, 용어해설 등을 점검하는 부분으로 나누어진다. 각 사안별로 정리한 결과는 다음과 같다.

(1) 수치자료

통계간행물과 통계 DB의 수치 일치 여부와 통계 개편 등으로 인한 작성 방법 변경이 공표 자료에 정확히 반영되었는지 점검한 결과 적절한 수준이었

다. 다만 통계표의 수치에 비상식적이거나 이상치 등은 발견되지 않았지만 가로합/세로합 불일치 건수가 상당 수 확인되었다.

(2) 통계표 형식 및 내용

통계표에 수록된 항목과 내용이 다른 부분이 몇 건 발견되었다. 다른 진단 항목에서는 문제가 발견되지 않았다.

(3) 용어 해설 부분

문제가 발견되지 않았다.

(4) 기타 오류

목차, 색인 등과 본문이 일치하지 않는 부분은 발견되지 않았으나, 몇 군데 단순한 오타가 발견되었다.

나. 이용자 편의사항 점검

(1) 이용자를 위하여

모든 진단항목에서 특별한 문제가 발견되지는 않았으나, 추후 마이크로데이터, DB 등을 공개하게 되면 제공 매체 제공 경로를 명시하는 것이 필요해 보인다.

(2) 조사 정보

대부분 진단항목에서 특별한 문제가 발견되지 않았다. 다만 공표 일정에 대한 예고가 없어 이용자들이 불편을 느낄 가능성이 있다.

(3) 모집단 및 표본 설계

목표모집단과 조사모집단이 불일치함에도 이에 대한 주의사항이나 비교에 대해 명시적으로 언급된 부분이 없어 이용자가 통계 수치를 오용할 가능성이 상당히 높다.

(4) 자료 집계 및 추정

주요 변수에 대해 상대표본오차를 계산해 수록했으나(2009 보고서 159쪽 이후) 이 변수들은 상대표본오차가 무의미한 변수들이다. 그리고 수록된 표본 오차값들은 실제로 허용오차에 해당하는 값인 것으로 판단된다. 기본적인 용어 사용에 주의가 필요하고 통계적으로 의미가 있는 수치에 한해 공표할 필요가 있다.

제 2 절 통계품질 개선 과제

앞 절의 진단 결과를 토대로 도출된 개선 과제는 다음과 같다.

1. 표본규모

표본추출틀과 표본규모에 관련한 문제이다. 서로 다른 기초조사에 근거한 두 개의 표본추출틀을 병용하기 때문에 공표된 수치의 의미가 불분명해지는 문제와, 표본규모가 전국 단위의 조사에 걸맞지 않게 너무 작아 공표된 값의 대표성에 문제가 생길 수 있다. 그러나 표본추출틀 문제는 본 통계의 목표모집단과 조사모집단을 일치시킬 수 있는 신규 기초 통계가 생기지 않는 한 현실적 해법이 존재하지 않는 것으로 보인다. 따라서 이 연구에서는 전국 단위 조사에 걸맞도록 표본규모를 늘리는 것을 장기 개선과제로 제시할 만하다. 적절한 표본규모의 수준을 결정하는 것은 통계 품질의 핵심이 될 만한

주요 항목을 선정하고 이에 해당하는 변수의 표본오차를 제어할 수 있도록 정하는 방식으로 이루어지면 된다. 또한 이제까지의 통계 자료에는 업종별, 규모별로 1차원적 통계표만 제시되고 있으나, 업종별*규모별의 2차원적 정보도 얻는 것이 가능하도록 표본 규모를 늘려야 한다.

2. 공표자료 정보량 제고

공표 자료의 정보량 손실을 줄여야 한다. 조사된 항목 중 미공표되거나 정보량이 상당히 손실된 채 공표되는 항목이 많다. 2009 보고서의 경우 조사표의 [PART 5] 『침해사고 피해』에서 조사된 항목의 공표에 특히 이런 경우가 많았다. 매출 손실 정도, 생산 효율 저하의 정도, 복구비용, 데이터 재생산 비용, 개인정보 유/노출 보상 비용 등이 조사되고 있으나 통계표를 살펴보면 이들 중 대부분의 항목에 대해 해당 침해사고의 경험 유무만을 따져 해당 빈도의 추정치만 자료를 공표하고 있어, 침해사고 피해의 유형별 심도(depth)에 대한 정보를 통계표로부터 얻을 수가 없다. 문제는 이러한 정보가 정보보호 관련 자료 이용자들에게 핵심적인 내용이라는 데 있다. 그러나 이러한 통계 수치의 신뢰도를 확보하려면 표본 규모를 늘리는 것이 선행되어야 한다. 그리고 침해 사고 발생 시 침해의 심도에 관련한 수치는 응답자가 정확히 파악하기 어려운 측면이 있고, 파악하고 있더라도 응답을 꺼리게 되는 경향이 있어 조사된 수치의 정확도가 매우 떨어지게 되는 문제가 있다. 따라서 원하는 정도의 신뢰도를 확보하려면 표본규모를 상당한 수준에까지 끌어 올려야 한다. 또한 침해사고에 의한 피해를 모든 조사 대상이 경험하는 것이 아니기 때문에 표본조사를 통해 모집단 전체의 침해의 심도를 추정하는 것은 매우 어려운 일이며, 보험계리학(Actuarial Statistics)에서 사용하는 확률통계적 모형을 활용하는 등 보다 심화된 연구가 필요한 것으로 판단된다. 이상을 종합할 때 중장기적 비전을 가지고 지속적으로 개선해 나가야 할 과제로 판단된다.

3. 시계열 자료 제공

시계열적 분석이 필요한 항목들에 대해 응답형식을 조속히 확정하고 DB를 체계적으로 관리할 필요가 있다. 매년 응답형식이 바뀌게 되면 시계열적 분석이 불가능해 생산된 통계의 정보 손실이 상당하다. 공표 자료 내에 장기적 시계열적 추세 등을 파악할 수 있도록 최소 10년 이상의 기간을 포괄하는 통계표를 제공한다면 이용자들의 활용도가 엄청나게 제고될 것으로 기대된다.

4. 공표 자료 포맷 다양화

공표 자료의 포맷을 다원화할 필요가 있다. 현행 공표 자료 외에 원자료에 대한 수요가 상당한 것으로 파악되었다. 온라인 상에서 DB 시스템을 통해 원하는 수준의 마이크로데이터에 접근 가능토록 시스템을 구축해야 한다. 또한 각종 통계표들을 이용자 편의에 맞는 포맷으로 내려받을 수 있도록 해야 한다. 현재는 pdf 형태의 보고서만 내려받게 되어 있는데 이를 text, MS-Excel, HWP, MS-Word 등 다양한 포맷으로 내려받을 수 있도록 해 이용자 편의를 제고할 필요가 있다.

5. 표본설계

본 조사의 표본설계에 과표본추출이 적용되었으나 사후추정을 위한 가중치(2009년 보고서 25쪽)에서 과표본추출을 적용한 흔적을 찾을 수 없다. 이는 사후추정에 의한 가중치를 부여했기 때문인 것으로 예상된다. 종사자 수 1,000명 이상인 사업체 혹은 셀의 모집단 크기가 10개 이하 경우에는 전수조사가 되도록 하는 것이 바람직하며, 사후추정을 위한 가중치의 특성상 어려움이 있는 경우에는 2개 이상은 되도록 해야 한다. 즉, 전수층의 기준을 명확하게 하는 것 역시 고려해야 할 사항이다.

사후추정을 위한 가중치의 경우 사업체 규모 250명 이상과 1,000명 이상을 통합하여 적용할 때 주의를 요한다. 특히 본 조사에서 적용하는 가중치

$$W_{gh} = \frac{N_{gh}}{n_{gh}} \cdot \frac{1}{R_{gh}}$$

를 적용할 때 전수조사에서 무응답이 있는 경우 가중치 적용 전에 가중치의 타당성에 대한 검토가 필요하다.

2009년 보고서 33쪽을 보면 표본오차 계산을 위해 SAS의 PROC SURVEYMEANS를 사용했다고 되어 있는데, 표본조사 표본설계의 표본오차 공식은 수식으로 제공하는 것이 바람직하다.

통계 설계 절차에서 당초 표본 설계대로 조사가 이루어질 수 있도록 해야 하고, 무응답이 발생한 경우 적절한 대체(imputation) 기법을 적용해 비표본 오차를 제어할 필요가 있다. 표본 설계 시 전수조사 층으로 설계되었음에도 실제 조사에서 무응답이 발생한 경우가 상당 수 있었다. 전수조사 층으로 설계되었을 때엔 해당 조사 대상들이 조사 결과에 매우 영향력이 큰 경우였을 것임을 고려하면 무응답을 적절히 대처하지 못하면 비표본오차가 엄청나게 증폭될 것이며 양적 변수의 경우는 공표자료의 대표성을 상실한다고 보아도 무방할 것이다. 따라서 추정값 산정시 가중치 혹은 대체에 따른 명확한 추정식의 제시가 필요하며 이는 전문가의 도움에 의해 이루어지는 것이 바람직하다.

6. 조사 위탁업체

통계작성 관련 예산의 대부분을 외부 조사기관에 조사를 위탁하는데 사용하고 있으나 납품받은 자료의 질적 수준이 다소 못미치는 것이 아쉬움으로 지적된다. 통계표 등의 질적 수준이 매우 낮으며 심지어 통계 패키지의 출력물을 그대로 옮겨 붙인 것으로 보이는 경우도 발견되었다. 작성기관 담당자가 자문진과 함께 책임있게 위탁기관의 납품물을 관리 감독할 필요가 있다. 표본설계내역을 반드시 보고서 형식으로 보관하고, 추정 및 마지막 표본오차

산정 단계까지 지속적으로 자문단의 자문을 받는 것이 바람직하다. 위탁기관의 조사 과정에도 적극적으로 개입할 필요가 있다. 작성기관 담당자가 조사기관 혹은 조사 현장에 직접 실사를 나가거나 지속적으로 조사 기관 업무 진행 상황을 파악해 관리해야 한다.

제 3 절 통계품질 개선 방안

앞 절의 개선 과제 논의 과정에서 대부분 자연스럽게 개선 방안이 함께 제시되었거나 직관적으로 개선 방안이 도출되었기에, 구체적으로 강조하여 다시 논의할 필요가 있는 과제만을 대상으로 논의하기로 한다. 아래에 논의하게 될 내용을 포함해 연구진에서 제안하는 구체적 개선 방안은 아래 <표 13>에 정리되어 있다.

1. 표본규모 및 표본설계

표본규모가 너무 작아 공표 자료의 통계표의 대부분이 업종별, 규모별로 각각 집계되었다. 그러나 업종별*규모별의 교차 테이블을 작성하는 것이 바람직하다. 이 때 40개의 셀(10개 업종과 4개 규모)의 교차 테이블을 작성해야 하는데 현재의 약 2,500개인 표본규모는 매우 부족한 수준이다. 따라서 가능한 교차테이블의 표본오차를 산정하여 국가승인통계의 기준에 맞는 타당한 표본규모의 산정이 필요하다. 다행히 2010년 조사는 표본 규모를 6,000~7,000 수준으로 늘려 진행하고 있으나 예산 사정으로 이 수준을 지속적으로 유지할 수 있을지 불투명하다. 일단 2010년 조사 결과를 바탕으로 교차표 형태의 자료 공표 방안에 대해 연구를 진행하는 것을 1차적 목표로 한다.

향후 2차원 교차 정보를 공표할 수 있는 표본 규모가 확보된다 하더라도 특정 정보보호 사안에 해당 사항(예: 침해 사고 경험)이 있는 경우에 한하여 조사가 유의미한 본 통계의 특성 상 통상적인 전국 단위의 조사보다 훨씬 더

큰 표본규모가 필요할 것으로 예상된다. 따라서 구체적인 표본규모 산정 및 표본설계는 정보보호와 표본설계에 모두 전문성을 갖춘 자문'단(團)'의 감수 하에서 이루어져야 한다. 조사를 마친 후 자료 분석 및 사후 관리까지 철저히 자문단의 감수를 받아야 한다.

이를 위해서는 결국 비용의 문제가 대두된다. 만약 예산을 당장 확보하는 것이 불가능하다면 단위 조사 비용 면에서 효율성을 확보할 수 있도록 조사표를 재설계할 필요가 있다. 현행 범주별 비율 위주의 공표 형태를 유지하는데엔 굳이 조사표 응답 형식을 구체적 수치를 적게 할 필요가 없다. 핵심적인 항목을 제외한 대부분 항목의 응답 형식을 범주형으로 조정한다면 피조사자의 부담이 줄고 단위별 조사비용을 절감할 수 있을 것이며 결과적으로 허락된 예산 하에서 어느 정도 수준은 표본규모를 늘릴 수 있는 여지가 생길 것으로 기대된다.

아울러 핵심 조사 항목으로 지정된 경우 최선을 다해 무응답 등으로 인한 자료의 결손이 발생하지 않도록 자료 품질을 관리하도록 해야 한다. 그리고 이들 항목에 대해서는 일단 교차표 형태가 아닌 전체 모집단 수준에서 구체적인 수치를 공표하기 시작하는 한편 장기적 시계열 형태로까지 공표 형태를 확대하는 것이 바람직하다.

2. 정보보호 수준 지표 개발

최근 스마트폰 보급 확대 및 무선 인터넷 환경의 확산으로 인해 정보보호 문제의 중요성이 날로 더해가고 있다. 본 통계의 조사 결과를 종합해 우리나라 전체의 정보보호 수준을 대표할 수 있는 지표(index)를 개발한다면 활용도가 매우 높은 정보가 될 것이며, 이 분야 통계에서 국제적으로 선도적 지위를 선점할 수 있는 수단이 될 것이다. 지수 산정 가정을 통해 조사항목의 중요도 및 정보보호지표로서의 유효성 등을 자연스럽게 검토하게 되어 본 통계의 품질을 발전시켜나가는 데 도움이 될 것으로 기대된다.

다만 급변하는 IT 환경에 적응성이 있는 평가 지표를 이용해 유효한 지수가 나올 수 있도록 구체적인 지수 산정 방식과 산업분야별 지수 산정 방식을 정하는 문제, 지수 산정을 위해 포괄하게 될 사업체 선정 등의 문제는 전문성을 가진 자문단의 도움을 받아 해결해야 한다.

<표 13> 통계품질 개선과제

개선과제	실행방법	기대효과	예상 문제점	비고
[중장기과제] 업종별*규모별 의 교차표 형태로 자료 공표	2차원 교차표 제공이 가능해질 수 있도록 표본규모 재산정을 위한 연구과제 추진	2차원으로 강화된 정보량을 활용해 국내 기업의 정보보호 실태에 대한 세밀한 분석 가능	본 통계 특성을 고려해 업종 및 규모 분류를 조정할 필요가 있음. 표본규모 확대로 예산 추가 확보 필요	관련: 10쪽, 46쪽
[단기과제] 장기적 시계열 자료 제공	주요 지표에 대해 시계열 자료 제시. 공표 대상 항목 응답형식 확정.	중장기적 추세 등의 파악이 가능해져 통계의 활용도 제고 가능	특별히 예상되는 문제점 없음	관련: 19쪽, 44쪽
[단기과제] 보고서의 영문요약본 발간	현행 보고서의 영문요약본 발간	국제적 벤치마크 및 국제적 표준을 선도하는 통계로 성장	특별히 예상되는 문제점 없음	관련: 53쪽
[단기과제] 이용매체 다양화	현행 pdf 형태로 제공되는 보고서 포맷을 다양화	이용자 편의가 획기적으로 개선되어 활용도 제고 가능	다양한 형태로 공표되는 자료의 일치성을 위해 세심한 관리 필요	관련: 11쪽, 44쪽, 51쪽
[단기과제] 공표절차 준수	공표 시점을 예고하고 준수	이용자 편의가 획기적으로 개선되어 활용도 제고 가능	특별히 예상되는 문제점 없음	관련: 26쪽
[중장기과제] 신규 이슈 문항 발굴 및 개발	시대적 흐름을 따라가는 이슈 문항 개발	적합성 측면 품질 제고	시계열 관리의 어려움	관련: 19쪽, 52쪽
[중단기과제] 정보보호 지표 개발	정보보호 수준 평가를 위한 지표 개발, 평가대상 사업체 선별, 지표 산정 방식 결정	정보보호실태를 요약한 유효한 정보 제공 가능. 통계 활용도 제고. 홍보에 활용 가능.	전문가 자문을 위한 예산 추가 확보 필요	관련: 47쪽, 52쪽

제 3 장 발전 전략 및 로드맵

제 1 절 발전 전략

1. 목표 개념

국내외 사회경제적 환경의 IT에 대한 의존도가 날로 높아져 가고 있으며, 앞으로 의존도가 심화되어 가는 속도가 더욱 빨라질 것으로 예상되는 실정이다. 또한 정보보호 관련 규제 및 국제적 표준이 확대될 것으로 예상되는 시대적 흐름을 고려할 때, 기업의 정보보호는 리스크(risk) 관리와 지속가능경영의 측면에서도 매우 중요한 요소이다. 이러한 환경에서 정보보호에 대한 정책적 관심과 실질적인 투자가 충분히 이루어져야 할 터인데, '민간기업정보보호실태조사'가 이를 위한 기초자료가 되어야 한다. 본 연구에서는 다음의 세 가지 목표 개념에 기반해 '민간기업정보보호실태조사'의 품질향상을 위한 발전전략을 제시하고자 한다.

첫째, **적확(的確, accurate)한 통계**이다. 이는 승인통계로서 의미를 가지기 위해 가장 기본이 되는 개념이다. 제공된 통계 자료의 수치의 신뢰성을 장담할 수 있어야 하며 시공간적 연속성을 유지해야 한다. 이는 통계품질의 6개 차원 중 '정확성', '비교성', '일관성'을 포괄하는 목표 개념이다.



<그림 8> 목표 개념

둘째, **적합(適合, adaptive)한 통계**이다. 급변하는 IT 산업 특성 및 이용자 수요에 대해 기민하게 대응하는 통계여야 한다는 의미에서의 적응성의 개념과, 조사 설계로부터 자료 분석 단계에 이르기까지 조사 대상의 특성을 고려해 실질적으로 작성 목적을 만족

시킬 수 있는 통계 자료를 생산해야 한다는 의미에서의 적합성의 개념을 포괄한다. 통계 품질의 6개 차원 중 '관련성'과 상통한다.

셋째, **유효(有效, effective)한 통계**이다. 이용자에게 실질적으로 의미가 있는 통계 서비스를 제공해야 한다는 목표 개념으로서, 통계 품질의 6개 차원 모두를 포괄하지만 '시의성/정시성', '접근성/명확성' 차원의 품질을 매우 중요한 요소로 하는 목표 개념이다.

2. 발전 과제

2.1 통계 자료의 이용자 접근수월성 확대

- 공표 형식을 한글, MS-WORD, MS-EXCEL 등의 포맷으로 다양화한다.
- 이용자 명부를 활용해 수시로 통계 관련 소식을 이용자들에게 RSS¹⁾ 형태로 제공한다.
- 온라인에서 제공하는 서비스를 포탈 개념으로 확장하여 이용자가 어느 수준까지는 온라인상에서 원자료를 직접 가공해 원하는 자료를 취득할 수 있는 시스템을 구축한다.

2.2 이용자 수요 파악 및 피드백 시스템 구축

- 통계 자료 중 활용도가 높은 항목과 그렇지 않은 항목을 파악한다.
- 공표범위를 세분화할 필요가 있는 항목을 파악한다.
- 조사를 원하는 지표 등을 이용자가 수시로 웹 상에 건의할 수 있는 창구를 마련하고 지속적으로 관리한다. 약간의 보상 체계를 마련한다.

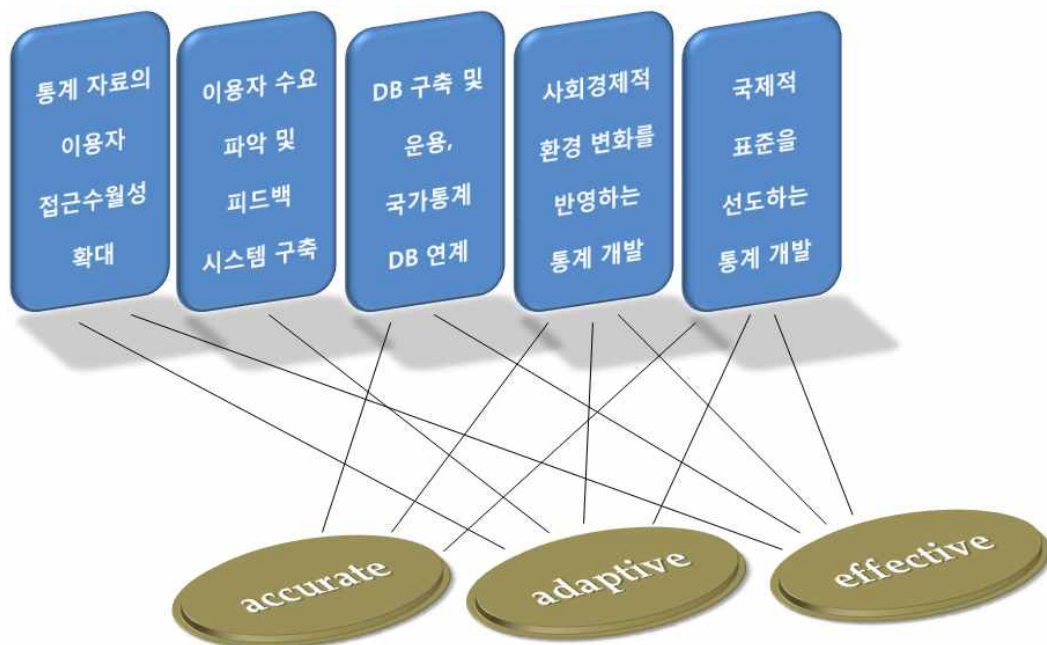
1) RSS는 'really simple syndication'의 머리글자를 딴 말로서, 업데이트되는 웹 사이트의 새로운 콘텐츠를 사이트 이용자들에게 제공함으로써 이용자가 가장 최근의 콘텐츠를 자동적으로 받아볼 수 있는 단일 소스를 갖게 하는 것을 뜻한다.

2.3 데이터베이스 구축 및 운용, 국가통계 DB와의 연계

- 데이터베이스 시스템을 구축하고, DB를 가공한 통계의 시공간적 연속성이 유지되도록 관리한다.
- 국가통계 DB에 연계하여 서비스를 제공한다.

2.4 사회경제적 환경 변화를 반영하는 통계 개발

- 정보보호의 수준이 기업 활동 성과에 직결되어 있으며, 기업의 지속가능 경영 수준과 유의한 상관관계가 있음을 널리 알린다.
- 시대를 이끄는 새로운 개념, 제품 등이 나왔을 때 빠른 시간 안에 이를 감지하여 통계에 반영한다.
- 사회경제적 환경 변화를 반영하여 우리나라 정보보호 수준을 가늠할 수 있는 지수(index)를 개발하고 지속적으로 관리한다. 업종별 특성을 고려한 업종별 지수를 함께 개발할 수 있다.



<그림 9> 5개 발전과제와 목표 개념

2.5 국제적 표준을 선도하는 통계 개발

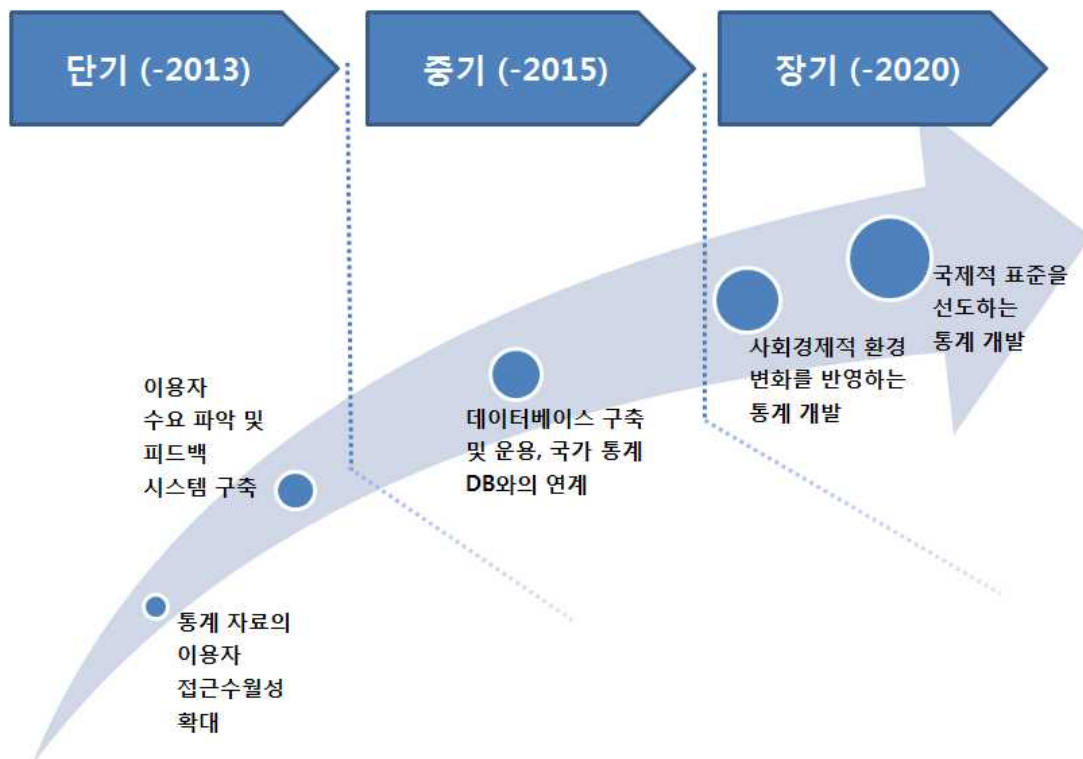
- IT 및 정보보안 관련 국제적 표준(global standard)이 아직 마련되어 있지 않다. 이 분야 최강국인 우리나라의 위상에 걸맞는 양질의 정보보호 관련 통계들을 생산하고 국제적 표준으로 제시할 필요가 있다. 현재 승인통계로서 이 분야에서 유일한 본 통계가 이러한 목표를 향해 나아가는 데 선도적인 역할을 해야 한다.
- 영문보고서를 발간하고 지속적으로 관리하는 한편 해외 유관 기관에 적극적으로 홍보한다.
- 조사항목 중 공표 시 누락되는 양적 변수(예: 침해사고 피해액)가 없도록 수치의 민감성을 고려하고 부정확한 응답을 방지할 수 있는 표본설계에 대한 연구를 진행한다.

제 2 절 로드맵

<그림 10>은 앞 절에서 제시한 5개 발전과제를 단기/중기/장기 과제로 나누어 간단히 로드맵의 형태로 제시한 것이다. 이 로드맵의 의의는 다음과 같다. 첫째, 통계 활용도를 높일 수 있는 이용자 중심의 계획이라는 점이다. 통계 품질에서 이용자의 만족도는 가장 중요하면서 1차적 요소라 할 수 있다. 그간 본 통계의 이용자가 주로 내부 이용자 및 정책이용자에 한정되어 있었으나, 이용자 수월성을 확대하고 이용자로부터 피드백을 받을 수 있는 장치를 마련함으로써 활용도가 일반 이용자들에게까지 획기적으로 확대될 수 있는 방안이 제시되었다. 둘째, 본 통계가 사회 경제적 변화에 발빠르게 대응하여 시의적절하게 유효한 정보를 제공함으로써 우리나라 정보보호 수준을 실질적으로 가늠해 볼 수 있는 통계를 생산하여 국가 경제에 이바지할 수 있는 방안이라는 점이다. 셋째, 아직 국내외적으로 표준이 제시되어 있지 않은 본 통계의 특성과 IT최강국으로서의 우리나라의 입지를 고려해 글로벌 스탠다드를 마련해 선도해 나가는 비전을 제시하고 있는 점이다. 소셜네트워

크 서비스(social network service, SNS)의 확대와 유비쿼터스(ubiquitous) 혹은 클라우드(cloud) 컴퓨팅 기술의 발전으로 정보보호에 대한 관심과 수요가 폭발적으로 늘어날 것으로 예상되는 바, 본 연구보고서에서 제시된 기본적인 문제점들부터 순차적으로 개선해 승인통계로서의 면모를 내실있게 확보한다면 그 자체로 이미 국제적 표준을 제시하는 일이 될 것이다.

따라서 이 로드맵에 기초하여 확실한 비전을 가지고 본 통계의 품질을 개선해 나간다면, 본 통계의 활용도를 더욱 높이고 국가 발전에 이바지하며 나아가 우리나라가 IT 강국으로서의 국제적 입지를 공고히 유지하는데 공헌할 수 있는 통계로 자리매김할 것으로 기대된다.



<그림 10> 5개 발전과제 및 로드맵

참고문헌

- [1] 방송통신위원회 한국인터넷진흥원 (2010). 2009년 정보보호실태조사 - 기업편.
- [2] 통계청 (2010). 통계품질이렇게합니다.
- [3] 통계청 통계개발원 (2007). 국가통계제도 비교를 위한 국제(UN) 공동 연구.
- [4] 통계청 통계교육원 (2010). 2010 교육훈련계획.

부록 1. FGI 결과보고 (전문이용자)



부	문	통계 1
통	계	명
승	인	번호
작	성	기관
F	G	I
일	자	
품질 진단 팀	연	구
	원	
	연	구
	보	조
	원	

제1부

회의 준비 및 진행과정

I. 회의 준비과정

1. 참석자 선정

<p>◦ 참석자 선정방법</p> <p>작성기관과 통계청에서 제공한 이용자 명부를 활용, FGI에 적합하다고 판단된 이용자를 선정</p>		<p>◦ 참석자 현황 (집단 구분에 ✓하고 인원 수를 각각 기입)</p> <p><input checked="" type="checkbox"/> 전문 이용자 집단</p> <p>- 정책고객(기업경영정책수립자) _____명</p> <p> 포함)</p> <p>- 교수 _____명</p> <p>- 연구원 _____명</p> <p>- 기타() _____명</p>
◦ 실시 장소	한국인터넷진흥원 서초청사 회의실	<p><input type="checkbox"/> 일반 이용자 집단</p> <p>- 대학원생 _____명</p> <p>- 대학생 _____명</p> <p>- 일반인 _____명</p> <p>- 기타() _____명</p>
◦ 소요 시간	120분	

2. 회의 참석자 명부

연 번	성 명	소속(부서명까지 기재)	직 위
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

II . 회의 진행과정

회의 진행	
<p>사전에 민간기업정보보호실태조사 보고서와 FGI질문지를 FGI 참석대상자에게 제공. 인터뷰 가이드는 품질차원별로 구분하여 제공.</p>	◦ 사회자 : 정석오
	◦ 기록자 : 김민영 김지학
	◦ 관찰자 :
	◦ 녹음·녹화 여부 : 녹음

제2부

회의록

○ FGI 실시과정에서 기록한 내용을 부문별로 정리

부 문	문제점 지적사항	개선의견 내용	핵심어
모집단 및 표본추출틀	목표모집단이 다른 사업체기초조사와 정보화 통계조사를 표본추출틀로 병용함	목표모집단과 조사모집단이 가능한 한 유사해질 수 있도록 노력	모집단, 표본추출틀
조사표 설계	조사표 문항이 너무 많음	문항에 따라 조사주기를 달리 하거나, 일부 조사 문항 삭제	조사표
공표 항목	조사된 항목 중 공표되지 않는 항목이 많음	충분한 표본 규모를 확보해 연속형 변수의 공표 수치의 신뢰도를 확보한 후 공표	연속형 변수
표본 규모	표본규모가 너무 작음	예산 추가 확보 및 효율화를 통해 표본 규모 증대	표본 규모
비교성	업종*규모 교차표 없음	표본 규모 확충을 통해 교차정보의 신뢰도 확보 후 공표	교차표
원자료 제공	원자료 제공하지 않음	DB 시스템 구축을 통해 제공	원자료, DB
공표 자료 형식	pdf로만 제공.	한글, MS-WORD, MS-EXCEL 등의 포맷으로도 제공	자료 형식
용어 사용	통계표에 부적절한 용어 사용 (예: 사례수)	수정	용어

※ 녹취록 및 질문지 별첨

제3부

FGI 결과 요약 및 개선 요구사항

○ FGI 회의록을 토대로 간략하게 정리

· 주요 토의 내용

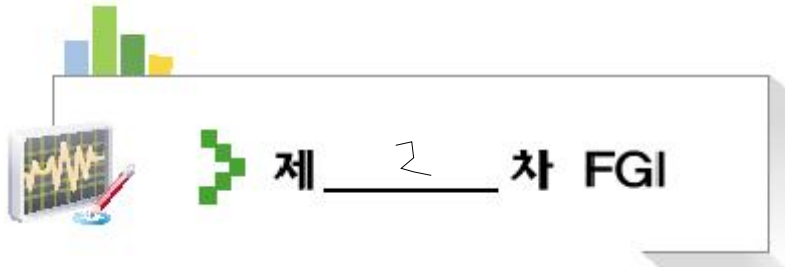
1. 목표 모집단과 조사모집단의 불일치 문제
2. 조사표 설계 시 조사항목의 수가 과도한 점
3. 연속형 변수에 해당하는 통계가 공표 자료에서 누락된 점
4. 표본규모의 문제
5. 통계표에 사용된 부적절한 용어 문제

○ FGI 회의록을 토대로 개선 요구사항 정리

· 주요 개선의견

1. 목표모집단과 조사모집단이 가능한 한 닮을 수 있도록 노력.
2. 조사표 내 조사항목별로 조사 주기를 조정하거나 조사항목 수 조정을 통해 피조사자 부담을 경감하고 수집자료의 신뢰도를 제고할 필요.
3. 표본규모를 확충해 연속형 변수 항목 및 업종*규모별 교차 정보의 제공이 가능케 해야.
4. 통계표 작성 시 면밀히 검토하여 부적절한 용어가 사용되는 일이 발생하지 않도록 해야.

부록 2. FGI 결과보고 (일반이용자)



부	문	통계 1
통	계	명
승	인	번호
작	성	기관
F	G	I
일	자	
품질진단팀	연구원	정석오
	연구보조원	김민영 김지학

제1부

회의 준비 및 진행과정

I. 회의 준비과정

1. 참석자 선정

<ul style="list-style-type: none"> 참석자 선정방법 <p>작성기관과 통계청에서 제공한 이용자 명부를 활용, FGI에 적합하다고 판단된 이용자를 선정</p>		<ul style="list-style-type: none"> 참석자 현황 (집단 구분에 <input checked="" type="checkbox"/>하고 인원 수를 각각 기입) <input type="checkbox"/> 전문 이용자 집단 <ul style="list-style-type: none"> 정책고객(기업경영정책수립자) _____명 포함 <ul style="list-style-type: none"> 교수 _____명 연구원 _____명 기타() _____명 <input checked="" type="checkbox"/> 일반 이용자 집단 <ul style="list-style-type: none"> 대학원생 _____명 대학생 _____명 일반인 _____명 기타() _____명
<ul style="list-style-type: none"> 실시 장소 	한국인터넷진흥원 서초청사 회의실	
<ul style="list-style-type: none"> 소요 시간 	120분	

2. 회의 참석자 명부

연 번	성 명	소속(부서명까지 기재)	직 위
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

II . 회의 진행과정

회의 진행	
<p>사전에 민간기업정보보호실태조사 보고서와 FGI 질문지를 FGI 참석대상자에게 제공. 인터뷰 가이드는 품질차원별로 구분하여 제공.</p>	◦ 사회자 : 정석오
	◦ 기록자 : 김민영 김지학
	◦ 관찰자 : 배수진
	◦ 녹음·녹화 여부 : 녹음

제2부

회의록

○ FGI 실시과정에서 기록한 내용을 부문별로 정리

부 문	문제점 지적사항	개선의견 내용	핵심어
업종 구분	이질적인 산업인 운수와 통신업을 묶어서 조사함	분리해서 조사 및 공표할 필요	업종 구분
표본설계 준수	전수조사 층에서 전수조사가 이루어지지 않음	표본설계대로 조사가 이루어질 수 있도록 관리	표본설계, 전수조사
원자료 제공	원자료를 제공하지 않음	DB시스템을 통해 원자료 제공 통로 마련	원자료, DB
조사항목	사안이 민감한 경우 조사가 실질적으로 불가능한 항목들이 존재	과감히 조사항목에서 제외. 공시 자료 활용 제고	fake answer

※ 녹취록 및 질문지 별첨

제3부

FGI 결과 요약 및 개선 요구사항

○ FGI 회의록을 토대로 간략하게 정리

· 주요 토의 내용

1. 본 통계의 특성에 맞게 업종 구분 조정 필요
2. 전수조사층으로 표본설계가 된 경우 설계 준수 필요
3. 원자료 제공 문제
4. 민감한 사안에 관련된 항목의 조사 결과의 신뢰도 문제

○ FGI 회의록을 토대로 개선 요구사항 정리

· 주요 개선의견

1. 운수업과 통신업을 분리
2. 전수조사층 설계 준수
3. DB시스템 구축을 통해 제한된 범위에서 원자료 제공
4. Fake answer를 피할 수 있는 장치 개발, 혹은 공시된 자료를 적절히 활용

부록 3. 수집자료 정확성진단 점검표

수집자료 정확성 점검 결과보고

부	문	통계1
통	계	명
승	인	번호
작	성	기관
품질진단팀	연구원	정석오
	연구보조원	김민영

제1부 점검계획

○ 점검을 위해 채택된 점검방법, 대상, 내용, 일정 등에 대하여 기술

1. 점검 방법

민간기업정보보호실태조사의 조사대행업체인 '한국갤럽'에서 부분별 정확성 점검에 필요한 인원을 모두 소집 후 관련 분야에 대한 인터뷰결과를 근거로 정확성 점검을 실시.

소집된 인원은 조사기획자 1인, 조사원관리자 3인, 조사원 3인과 응답자 3인.

1. 표본조사에 대한 총체적, 개괄적 점검 - 조사기획자.

- ① 자료수집 방법의 적절성
- ② 현장점검 및 관리체계
- ③ 대상처/응답자 관리

2. 조사원들의 역할에 대한 관리 - 조사원관리자

- ① 조사원 관리
- ② 대상처/ 응답자 관리

3. 실제 조사원들의 업무 실태와 현장 관리 - 조사원

- ① 업무량 및 교육
- ② 대상처 / 응답자 관리

2. 면담(현장방문) 일정

일시	면담대상자/참석자	장소	주요 점검사항

제2부 점검결과 요약

○ 점검결과 주요 문제점 및 개선의견 정리

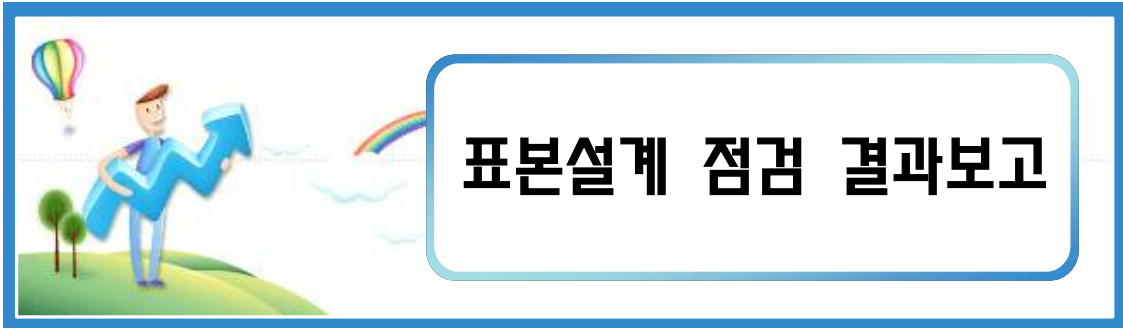
부 문	문제점	개선의견
[조사기획자] 자료수집방법	무응답 처리에 대해 고려하지 않음	무응답 대체를 고려할 필요가 있음
[조사기획자] 현장점검 및 관리체계	표본설계, 실사관리원 관리, 통계처리, 조사결과 분석 등 조사업무 전반에 대한 관리 및 감독을 담당하고 있으며, 실사관리원과의 현장 점검 등을 통한 즉각적인 대처, 조사결과의 시계열적 확인 등의 조사 전반에 대한 관리 및 점검이 필요하므로 업무량이 많은 편	인력 확보 및 전문화를 통해 업무량을 분산시킬 필요가 있음
[조사원관리자] 대상처/응답자 관리	응답거부율이 높음.	조사의 중요성을 잘 설명하는 한편, 선물의 수준을 높일 필요가 있음
[조사원] 업무량 및 교육	조사 특성 상 업무량이 많은데 비해 임금이 낮게 책정되어 있음	철저한 사전 교육을 통해 조사원들이 업무에 익숙해질 수 있도록 하여 효율적으로 업무가 처리될 수 있게 함

제3부 점검결과 종합

- 점검결과를 통해 현장조사의 오류 유형과 발생 원인을 종합적으로 분석하고, 정확성 제고를 위한 방안 기술

조사 전문 기관인 한국갤럽이 조사를 진행한 때문인지 전반적으로 자료 수집 절차는 양호한 품질을 가지고 있는 것으로 진단되었다. 다만 낮은 응답 거부율을 제고할 수 있는 방안으로 본 통계의 중요성에 대해 적극적으로 홍보를 실시하는 한편 응답자 선물의 수준을 현격히 높이고 필요 시 공문 발송 등을 고려할 필요가 있다. 또한 작성기관 담당자가 조사 기관 혹은 조사 현장에 직접 실사를 나가거나 지속적으로 조사 기관 업무 진행 상황을 파악해 관리함으로써 납품물의 품질을 관리할 필요가 있다.

부록 4. 표본설계 진단 결과



부	문	통계 1
통	계	명
승	인	번호
제	34201	호
작	성	기관
한	국	인터넷진흥원
품질 진단 팀	연구원	정 석 오
	연구보조원	김 민 영

점검 개요

○ 표본설계 점검 시 검토한 자료(표본보고서 등), 면담자, 면담일시 등 기술

일시 : 2010년 7월 14일

장소 : 경기대학교

면담 및 참석자 : 경기대 통계학과 이상은 교수
 한국외국어대 통계학과 신기일 교수
 한신대 통계학과 한근식 교수
 우석대 통계학과 이기성 교수

검토자료 : 표본설계 요약 보고서 및 민간기업정보보호실태조사 보고서

조사 개요

조 사 명	민간기업정보보호실태조사	
작성기관명	한국인터넷진흥원	
전 수 / 표 본 조 사	전수()	표본(0)
표 본 설 계 주 체	자체설계()	외부용역(0) 【용역사업자: 】
조 사 목 적	국내 기업 및 기관의 정보보호실태 및 문제점에 대한 현황 파악과 원인 분석을 통해 정보보호 인식 제고 및 수준 향상을 위한 정책 개발	
조 사 대 상	전국 종사자수 5인 이상, 네트워크로 연결된 PC를 1대 이상 보유한 사업체 2,234개	
조 사 방 법	방문면접조사	

표본설계 개요

구분	내용
모집단	종사자수가 5인 이상이고 네트워크가 구축된 컴퓨터가 1대 이상인 사업체
표본추출틀	1. 사업체기초통계조사에서 대상 사업체 2. 2009년 정보화통계조사 사업체 중 네트워크 구축 사업체 : 구축률을 활용 => 한국표준산업분류의 의한 20개의 대분류 중 가사서비스업, 국제 및 외국기관, 자동차관련도·소매업을 제외한 18개 산업에 속하고 종사자 수가 5명 이상인 사업체 531,345,개 중 네트워크에 연결된 컴퓨터를 1대 이상 보유한 사업체 308,529개 업체
표본추출방법	<ul style="list-style-type: none"> · 2008년 사업체기초통계조사 자료에서 5인 이상 사업체들을 업종별(10종)× 규모별(5종)로 총 50개의 층으로 분류 · 2009년 정보화통계조사 자료를 이용하여 모집단의 분포를 파악 · 예산에 의해 정해진 약 2,500개의 전체 표본규모를 업종별 규모에 따라 비례 배분 · 업종별과 규모별로 2단 층화한 후 각 사업체들을 지역별로 정렬하여 계통추출을 기본으로 함 · 절사추출 : 종사자 수가 1,000명 이상인 사업체는 전수조사 실시, 종사자 수 250~99명인 사업체는 일부 전수조사. · 과표본추출 : 50개 층별로 최소 20~30개 우선 할당. · 비례할당 : 업종 및 규모별 조사 가능성을 고려하여 배분.
표본크기	2,234
가중치	<ul style="list-style-type: none"> · 본 조사는 층별 비례할당을 하지 않았기에 모집단과 표본간의 편차 최소화를 위해 가중치를 이용한 사후층화를 이용함. · 비율이 낮은 250이상 규모의 사업체는 통합 계산.

	<ul style="list-style-type: none"> · 가중치 계산 $W_{gh} = \frac{N_{gh}}{n_{gh}} \cdot \frac{1}{R_{gh}}$ N_{gh} = g업종의 h규모의 조사모집단의 크기 n_{gh} = g업종의 h규모의 층에 할당된 표본크기 R_{gh} = g업종의 h규모의 층의 응답률
추정식	<ul style="list-style-type: none"> · 허용오차 $e = Z_{\alpha/2} \times \sqrt{P(1-P)/n(1-n/N)}$ N = 모집단 크기 n = 표본 크기 · 모수 추정 $\hat{Y} = \sum_{h=1}^L \hat{Y}_{hc} + \sum_{h=1}^L \frac{N_{hs}}{n_{hs}} \sum_{k=1}^{n_{hs}} y_{hsk}$ L = 층의 개수 (업종 \times 규모) n_{hs} = h층의 표본 크기 y_{hsk} = h층 표본의 k번째 관측값 \hat{Y}_{hc} = 전수조사에서 각 층의 총계에 대한 추정량

점검결과 요약

○ 점검결과 주요 문제점 및 개선의견 정리

부문	문제점	개선의견
표본규모 문제	표본규모는 주요 항목에 대한 표본오차를 기준으로 정해지는데 본 통계는 '정보화 투자 대비 정보보호 투자 비율'항목을 기준으로 전국 단위의 표본오차를 구하였다. '지출없음'과 '1%미만'의 셀은 표본오차(RSE)가 5% 미만으로 나타났으나, 전체적으로는 5%를 훨씬 초과하고 있는 것으로 나타났다.	다양한 항목을 동시에 고려해 표본오차를 검토하고 적절한 수준으로 이들을 제어할 수 있도록 표본규모를 늘릴 필요가 있다.
부적절한 용어 사용	본 조사의 제목은 '민간기업정보호실태조사'이다. 따라서 민간기업을 기준으로 조사 결과를 보여주는 것이 바람직하다. 그러나 모집단과 표본들을 살펴 보면 표본추출단위가 사업체 단위로 되어 있어 모든 조사의 집계 결과는 사업체 단위로 이루어진다. 따라서 '민간기업..'이라기보다는 '민간사업체..'로 보는 것이 바람직하다. 표본설계보고서에서도 정확한 용어의 사용이 필요하다. 예를 들면 허용오차와 표본오차의 용어가 혼	적절한 용어로 대체가 필요하다.

	용 내지는 오용되고 있다. 그리고 통계결과표는 총계 추정치를 제시하고 있는데 이 때 총계 추정치를 사례수로 표기하는 것은 적절치 않다.	
표본설계보고서 문제	표본설계는 외부용역에 의해 이루어지고 있으며 표본설계내역을 별도의 보고서 형태로 받지 않고 대부분 표본리스트와 간단한 설계내역만을 받고 있다. 이렇게 되면 담당자들의 표본설계 및 추정에 관한 내용에 대한 이해도는 피상적일 수밖에 없으며, 추정 과정 및 표본오차의 검토 단계가 소홀해지게 된다.	표본설계내역을 보고서 형식으로 보관하고, 추정 및 마지막 표본오차 단계까지 지속적인 자문이 이루어지는 것이 바람직하다.
통계표 일치성 문제	통계표의 기준이 같을 경우에는 통계표들의 전체 합계의 값은 같아야 하는 것이 일반적이거나, 본 조사의 보고서는 전체 값이 일정하지 않다.	무응답 및 결측값 처리를 고려함으로써 해결 가능할 것으로 보인다.

점검결과 종합

- 점검결과를 종합적으로 분석하여 현재 표본설계 상 보완이 필요한 사항, 개선방안, 발전전략 등 제시

- 본 조사의 표본설계내용에 과표본추출이 적용되었다. 그러나 사후추정을 위한 가중치(보고서 25쪽)에서는 과표본추출이 적용을 볼 수가 없다. 이는 사후추정에 의한 가중치 부여에 따른 것으로 예측된다, 그러나 1000명 이상인 사업체 혹은 셀의 모집단이 10개 이하 경우에는 전수조사가 되도록 하는 것이 바람직하나 사후추정을 위한 가중치의 특성상 어려움이 있는 경우에는 2개 이상은 되도록 하는 것이 바람직하다. 즉, 전수층의 기준을 명확하게 하는 것 역시 고려해야 할 사항이다.

- 사후추정을 위한 가중치의 경우 250명 이상과 1000명 이상을 통합하여 적용할 때 주의를 요하게 된다. 특히 본 조사에서 적용하는 가중치, $W_{gh} = \frac{N_{gh}}{n_{gh}} \cdot \frac{1}{R_{gh}}$ 를 적용할 때 전수조사에서 무응답이 있는 경우 가중치 적용 전 가중치의 타당성 검토가 필요하다,

- 표본조사의 표본설계에 표본오차 공식은 수식으로 제공하는 것이 바람직하다.

- 전수조사부분에서는 추정량이란 용어는 어색한 듯 하다.

- 통계결과표는 총계 추정치를 활용하였다. 이 때 총계추치를 사례수로 표기하는 것은 어색한듯하다. 또한 통계표의 기준이 같을 경우에는 통계표들의 전체 값은 같아야하는 것이 일반적이나 본 조사의 보고서는 전체 값이 일정하지 않다. 이는 무응답 및 결측값 처리를 전혀 고려하지 않았기 때문으로 보인다.

- 2009년 보고서의 343쪽에서 ~361쪽의 자료검토가 철실하게 필요부분이다.
- 표본오차를 실어주는 것이 바람직하다.
- 조사표에서 범주형자료 보다는 가능한 연속형자료로 얻어지는 것이 통계의 활용도에서 매우 좋다. 특히 본 조사에서 종사자수는 범주를 정해서 얻는 것보다는 종사자수를 조사하는 것이 바람직하다. 또한 피해규모는 조사가 이루어지나 통계 결과표에는 활용되지 않고 있다. 조사가 이루어진 부분의 통계표 작성에 활용하는 것이 바람직하다.
- 통계표 한 변수의 분류에 의해 작성되고 있다. 최소 2개의 변수에 의한 테이블 작성도 필요한 듯하다. 즉 cross table 작성도 고려해 보아야 할 듯하다.
- 표본오차에 의한 표본규모 산정도 시도해 볼 만하다.