

정책보고서 2009-04

2009년 전국 보육실태조사 표본설계

Sampling Design for 2009 National Child Care Actual Conditions Survey

손 창 균

한 국 보 건 사 회 연 구 원
보 건 복 지 가 족 부

제 출 문

보건복지가족부 장관 귀하

본 보고서를 보건복지가족부의 『2009년 전국보육실태 조사 표본설계』에 관한 연구용역 최종보고서로 제출합니다.

2009년 2월
한국보건사회연구원
원 장 김 용 하

머리말

저출산 고령화 사회에 본격적으로 접어든 현재 급격한 출산을 저하에 따른 각종 아동관련 복지욕구 및 수요의 변화에 빠르게 대처하기 위해 2005년에 발효된 영유아 보육법에 따라 매 5년마다 실태조사를 통해 급격히 변화하고 있는 자녀의 양육 환경에 부응하여 각종 영유아 관련 정책수립의 준거가 될 수 있는 객관성 있는 기초 자료의 생산을 목적으로 『전국보육실태조사』를 실시하고 있다.

2004년에 실시된 조사에서는 보육 및 교육이용 관련 실태조사를 실시하였고, 여성부, 교육부 보건복지부 등의 3개 정부부처가 공동으로 사업을 수행하였다. 2004년 조사에서의 주요 특징을 살펴보면 다음과 같다.

첫째, 전국을 5개 층(서울의 동, 광역시의 동, 기타시의 동, 읍, 면)으로 구분하여 표본설계가 이루어졌으며, 조사 모집단은 2000년 인구주택 총조사 10% 표본자료를 이용하였다. 한편 전체 표본수는 12,000가구를 목표 표본수로 하였고, 200개의 표본조사구를 추출하였다.

둘째, 2004년 표본설계에서 고려한 표본 배분 방식이나, 표본추출방식에 관한 사항이 언급되지 않았으나, 통상적으로 조사구 특성별로 정렬한 후 확률비례 계통 추출을 사용하였고, 별도의 표본배분 방식을 이용하기 보다는 3개 층에 대한 조사구 비율에 따라 적정한 추출율을 적용하여 최종적으로 표본조사구를 추출한 것으로 판단된다.

이와 달리 2009년 조사의 경우 조사모집단은 2005년 인구주택 총조사 90%자료를 활용하여 표본추출이 이루어졌으며, 2004년과 달리 전국을 3개 층(대도시, 중소도시, 읍면)으로 층화하고, 각 층별 표본배분은 층별로 0~6세 이하의 아동수에 비례하여 표본을 배분하도록 하였다.

전체 표본수에서는 아동출현율을 감안하여 2004년 보다 약 50개 조사구를 확대하여 총 257개 표본조사구를 추출하여 조사구당 60가구를 조사하도록 하였고, 목표 표

본수는 15,000가구로 결정하였다.

최근 급격한 출산 수준의 감소는 자녀양육지원 정책을 강화하는 배경이 되고 있는데, 2005년 1.08명, 2006년 1.13명, 2007년 1.26명으로 2005년 이후 점차적으로 개선되고는 있지만, 매우 낮은 수준으로 이와 같은 낮은 출산율의 원인으로는 여성의 사회참여확대, 높은 양육부담, 자녀양육에 대한 여성의 부담, 자녀에 대한 노후 보장 기대의 감소, 결혼가치관의 변화 이 매우 큰 원인으로 파악되고 있기 때문에 이에 대한 적절한 정책 수립이 필요한 실정이다.

이와 같이 육아보육의 부담이 갈수록 증가되는 상황에서 정부의 육아보육실태에 대한 개관적인 통계자료의 산출을 통해 육아보육 수요자의 욕구 및 만족도를 파악하여 육아보육정책의 올바른 방향성을 설정하고, 육아보육 지원 예산의 합리적 집행을 통한 국가 재정의 효율화에 기여할 수 있을 것으로 사료된다.

본 연구는 손창균 부연구위원의 책임하에 수행되었으며, 바쁜 연구중에도 본 보고서를 검독해주신 장영식 연구위원과 도세록 연구위원, 연구자문을 해주신 홍기학 교수와 이기성 교수께도 감사를 드리며, 보고서 편집에 도움을 준 김은주 연구원께도 감사를 드린다.

끝으로 본 연구에 수록된 내용은 어디까지나 연구자의 의견이며, 본 연구원의 공식적인 견해가 아님을 밝혀둔다.

2009년 2월

한국보건사회연구원

원 장 김 용 하

목 차

ABSTRACT	11
요 약	13
제1장 서론	19
제1절 연구 목적 및 배경	19
제2절 연구내용 및 방법	21
제3절 표본설계의 기본 목표와 절차	24
제2장 2004년 조사에 대한 분석	25
제1절 2004년 조사의 개요	25
제2절 2004년 조사의 특징	28
제3절 2004년 표본설계 및 특성별 분석	29
제3장 2009년 전국 보육실태조사 표본설계	48
제1절 2009년 표본설계의 특징	48
제2절 모집단 분석	49
제3절 층화(Stratification)	61
제4절 표본 규모의 결정 및 배분	64
제5절 표본 추출방법	71

제4장 추정	73
제1절 기본 가중치 산정	73
제2절 특성별 추정	74
제5장 가중치 조정방법	78
제1절 기본가중치	79
제2절 무응답 조정가중치	79
제3절 사후층화 가중치	81
제4절 가중치의 효과	83
제5절 가중치의 절단	84
제6절 가중치 조정시 고려 사항	86
제6장 조사 및 비표본오차의 관리	88
제1절 조사관리	88
제2절 비표본오차 관리	90
제7장 제언	94
참고문헌	96
부 록	101

표 목 차

〈표 2- 1〉 2004년 조사의 표본 조사구 특성	27
〈표 2- 2〉 2004년 조사결과의 지역별 특성	29
〈표 2- 3〉 16개 광역시도별 조사대상 아동 수에 대한 지역별 변동계수	30
〈표 2- 4〉 5개 권역별 조사대상 아동 수에 대한 변동계수	30
〈표 2- 5〉 3개 권역별 조사대상 아동 수에 대한 변동계수	31
〈표 2- 6〉 층화기준에 따른 주택소유형태별 아동출현율 추정치와 변동계수	31
〈표 2- 7〉 16개 광역시도별 주택소유 비율에 따른 아동출현율에 대한 변동계수	32
〈표 2- 8〉 5개 권역별 주택소유형태에 따른 아동출현율에 대한 변동계수	35
〈표 2- 9〉 3개 권역별 주택소유형태에 따른 아동출현율에 대한 변동계수	37
〈표 2-10〉 층화기준에 따른 주택유형별 아동출현율에 추정치와 변동계수	38
〈표 2-11〉 16개 광역시도별 주택유형에 따른 아동출현율에 대한 변동계수	39
〈표 2-12〉 5개 권역별 주택유형에 따른 아동출현율에 대한 변동계수	42
〈표 2-13〉 3개 권역별 주택유형에 따른 아동출현율에 대한 변동계수	43
〈표 2-14〉 지역별 아동 연령대별 분포	44
〈표 2-15〉 5개 권역별 아동 연령별 분포	45
〈표 2-16〉 3개 권역별 아동 연령별 분포	45
〈표 3- 1〉 조사구 특성별 분포	48
〈표 3- 2〉 거주 종류별 가구수	48
〈표 3- 3〉 지역별 조사구, 가구 및 인구 분포	49
〈표 3- 4〉 지역별 조사구 분포(아파트, 일반)	50
〈표 3- 5〉 지역별 가구 분포(아파트, 일반)	51
〈표 3- 6〉 지역별 인구 분포(아파트, 일반)	52
〈표 3- 7〉 지역별/조사구유형별 만 0~6세 아동 인구분포(90%자료)	53

〈표 3- 8〉 지역별/조사구유형별 0~6세 아동 연령의 CV분포(90%자료)	54
〈표 3- 9〉 5개 권역별/조사구유형별 0~6세 아동 수 및 CV분포(90%자료)	55
〈표 3-10〉 3개 권역별/조사구유형별 0~6세 아동 수 및 CV분포(90%자료)	56
〈표 3-11〉 지역별, 성별 만 0~4세 추계 인구분포	57
〈표 3-12〉 지역별, 설립주체별 보육시설 현황(2008년 6월말 현재)	58
〈표 3-13〉 지역별, 도시규모별 보육시설 현황(2008년 6월말 현재)	59
〈표 3-14〉 95% 신뢰수준 하에서 허용오차와 보육 수요율에 따른 표본수	63
〈표 3-15〉 5개 권역별 조사구 규모에 따른 층별 비례배분 현황	65
〈표 3-16〉 5개 권역별 네이만 배분에 의한 층별 표본 배분 현황	65
〈표 3-17〉 3개 권역별 비례배분에 의한 층별 표본 배분 현황	66
〈표 3-18〉 3개 권역별 네이만 배분에 의한 층별 표본 배분 현황	66
〈표 3-19〉 5개 권역별 아동수에 따른 층별 비례배분 현황	67
〈표 3-20〉 5개 권역별 아동 연령 변동에 따른 층별 네이만 배분 현황	68
〈표 3-21〉 3개 권역별 비례배분에 의한 층별 표본 배분 현황	68
〈표 3-22〉 3개 권역별 네이만 배분에 의한 층별 표본 배분 현황	69
〈표 4- 1〉 추출설계와 가중치	77
〈표 5- 1〉 주요 국가의 가계조사 무응답률	87
〈표 5- 2〉 통계청 경상표본조사에서의 당일 방문 면접완료율	87

그림 목차

〔그림 3-1〕 2009년 보육실태조사 표본설계 절차	50
〔그림 6-1〕 조사연구과정에서 발생하는 오차	93

ABSTRACT

Sampling Design for 2009 National Child Care Actual Conditions Survey

Recently the rapidly decrease of child birth rate becomes the background of strengthen for child care supporting policy. Even though the total fertility rate was gradually changed after 2005, such as 1.08 person in 2005, 1.13 person in 2006, and 1.26 person in 2007, the child birth rates are very low level compared that of OECD. The main reasons of low level of child birth rate are known to the expansion of women's social participation, the high child rearing burden, the decreasing expectation of the security of aged, and the changing of marriage standard so that it needs to establish the suitable policy for the low fertility.

To know this present state of low child birth rate, the several government agencies included KIHASA conducted the national survey for child care and education actual conditions in 2004, where it was drawn by PPS 12,000 households in 200 ED's from the Population and Housing Census in 2000 which implemented by KNSO.

Now, we suggest the difference sampling schemes compared to 2004 in terms of the target population, sampling method, and sample allocation.

In this sampling design for 2009 National Child care Actual Conditions Survey it is designed the 3 strata such as "Large Cities", "Medium and Small Cities", and "Agricultural and Fishery Areas(Eup and Myeon)". The sampled PSU's are drawn 257

ED's and SSU's 15,000 households by PPS from the 2005 Population and Housing Census 90% data. Also, the sample allocation is used the proportional allocation according to the numbers of 0~6 aged child by strata. We also propose the weighting adjustment and the management of nonsampling errors occurred in survey procedure.

요 약

제1장 서 론

제1절 연구배경 및 목적

□ 연구배경

- 『전국보육실태조사』는 한국보건사회연구원에서 제한된 환경에서 2001년 전국 표본 100개 지역 약 12,000가구에 거주하는 초등학교 이하 아동의 보육 및 교육서비스 이용실태를 파악하여 연령별, 지역별 보육수요율을 산출하여 중기 영유아 보육수요율을 34.1%로 산정하는 등의 연구결과를 제시하여 보건복지부의 보육정책을 위한 기초 자료로 활용한 바 있음.
- 2004년 보육수요 및 교육이용실태 조사에서는 한국보건사회연구원을 포함한 여러 기관의 공동연구로 전국 200여개 조사구, 12,000여 가구를 조사하여 최종 10,423가구에 대해 조사 완료하여 87.1%의 완료율을 보였고, 이중 6,412명의 아동을 대상으로 조사완료 되었음.

□ 연구목적

- 『전국 보육실태조사』는 국가승인통계(승인번호: 제33107호)로서 2005년에 발효된 영유아 보육법에 따라 매 5년마다 실태조사를 통해 급격히 변화하고 있는 자녀의 양육환경에 부응하여 각종 영유아 관련 정책수립의 준거가 될 수 있는 객관성 있는 기초 자료의 생산에 목적을 두고 있음.
- 이를 위해 수요자 지향적인 보육 서비스 공급을 위한 자료를 산출하여 지역, 아동연령, 소득수준, 母의 취업 등 가정 및 아동특성별 보육서비스 이용 및 수용자의 욕구를 파악하고, 수용자 특성, 요구 및 만족도를 파악함으로써 제도개

선을 위한 정책 자료를 생산함.

- 각종 육아정책수립을 위한 정부 공식통계 자료를 확보하고, 향후 보육정책 수립 및 각종 육아정책을 위한 공식통계를 제공함으로써 정부 정책의 타당성과 일관성을 보장하기 위함.

제2절 연구내용 및 방법

- 다음과 같은 내용으로 연구를 수행하도록 함.

- 과거 조사 자료에 대한 분석
- 모집단 분석
- 2009년 표본설계
- 추정식 유도
- 가중치조정
- 비표본오차관리

제3절 표본설계의 기본 목표와 절차

- 전국을 대도시, 중소도시, 농어촌으로 층화하여 영유아의 보육수요율을 정확히 파악할 수 있도록 표본 설계함.
- 조사연구의 목적에 가장 부합되는 조사모집단을 설정하고 최선의 추출틀을 사용하여 표본이 모집단의 특성을 잘 대표할 수 있도록 함.
- 과거 표본설계의 장·단점을 분석하여 문제점을 보완해서 조사결과의 정확성과 신뢰성을 높이도록 함.
- 확률추출을 근간으로 해서 전국단위의 보육수요율 및 대도시, 중소도시, 농어촌별 추정이 가능하도록 추정량과 추정량의 분산 추정식을 유도함.
- 조사 후 추정 시 문제가 되는 가중치 조정방법 및 과거 조사과정을 검토해서 조사의 효율성을 높이는 방안 등을 고려함.

제2장 2004년 조사에 대한 분석

제1절 조사개요 및 방법

- 조사원에 의한 가구방문 면접조사로 수행되었고, 가구조사표와 함께 아동이 있는 가구에 대해서는 개별아동조사표를 함께 조사하도록 하였음.
- 조사표는 여성부, 교육인적자원부, 대통령자문 고령화 및 미래사회위원회 정책담당자, 보육시설의 장으로 구성된 자문단에 의해 검토되었음.

제2절 2004년 조사의 특징

- 전국 12,000가구를 표본추출하여 가구원 및 가구의 특성을 파악하여 초등학교 이하 아동이 있는 경우 이들에 대한 개별조사를 수행함.
- 조사모집단은 2000년 인구주택총조사 10% 조사구 자료이며, 3개 층으로 층화하여 각 층의 주택 유형에 따라 정렬한 후 200개 조사구를 표본추출함.
- 최종 조사 완료된 가구 수는 10,423가구 아동은 6,412명이 조사 완료됨.

제3절 2004년 표본설계 및 특성별 분석

- 첫째, 표본 추출틀로 2000년 인구주택 총조사의 10% 조사구 자료를 이용하였음.
- 둘째, 조사구 특성별로 정렬한 후 확률비례 계통 추출을 사용하였고, 3개 층에 대한 조사구 비율에 따라 적정한 추출율을 적용하여 최종적으로 표본조사구를 추출함.
- 셋째, 별도의 표본 규모에 대한 산정을 하지 않음.
- 넷째, 2000년 인구주택총조사 10%자료를 이용하여 표본추출이 되었기 때문에 추출틀의 노후화에 따른 별도의 부가적인 추출틀을 활용하여 표본을 추출함.

제3장 2009년 표본설계

- 2009년 전국보육실태조사의 조사 모집단은 2005년 인구주택총조사 90%자료임.
- 전국을 3개층(대도시, 중소도시, 읍면)으로 층화하여 각 층별로 0~6세 이하의 아동 수에 비례하도록 표본을 배정함.
- 표본 PSU는 257개 조사구이며, 이는 보육수요율을 40%로 가정하여 95% 신뢰수준하에서 허용오차 약 6% 이내로 오차를 관리할 수 있는 수준이며, 최종적인 표본은 PSU당 평균 60가구를 가정할 경우 총 15,420가구를 표본추출하게 됨.
- 층화는 5개 권역과 3개 권역을 비교하여 최종적으로 각 층별 보육수요율이 안정적으로 추정될 수 있도록 하기위해 3개 권역별 층화로 결정함.
- 최종적으로 결정된 층별 표본수는 3개 층별 0~6세 이하의 아동수에 비례하여 배분하여 대도시의 일반조사구 50개, 아파트조사구 60개, 중소도시 일반조사구 35개, 아파트조사구 68개, 읍면지역의 일반조사구 20개, 아파트조사구 24개로 배분되었음.

제4장 추정

제1절 기본 가중치 산정

- 각 층별로 배분된 조사구 규모에 따라 모집단 층으로부터 추출된 단위들의 추출확률(p_{hij} : sampling probability)는 다음과 같음.

$$p_{hij} = p_{hi} \times p_{hj(i)} = \frac{n_h}{N_h} \times \frac{m_h}{M_{hi}} = \frac{n_h m_h}{N_h} \times \frac{1}{M_{hi}}$$

- 따라서 비례배분 하에서 표본으로 추출된 가구의 가중치(w_{hij} : sampling weight)는 다음과 같음.

$$w_{hij} = \frac{1}{p_{hij}} = \frac{N_h}{n_h m_h} \times M_{hi}$$

제2절 특성별 추정

- 전국 아동 수 총계, 분산 및 상대표준오차 추정

$$\hat{Y} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}, \quad var(\hat{Y}) = \sum_{h=1}^L \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi.} - \overline{y_{h..}})^2,$$

$$\widehat{CV}(\hat{Y}) = \frac{\sqrt{var(\hat{Y})}}{\hat{Y}} \times 100 (\%)$$

- 각 특성별 아동 수 총계, 분산 및 상대표준오차 추정

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}, \quad var(\hat{Y}_h) = \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi.} - \overline{y_{h..}})^2,$$

$$\widehat{CV}(\hat{Y}_h) = \frac{\sqrt{var(\hat{Y}_h)}}{\hat{Y}_h} \times 100 (\%)$$

- 전국 아동의 보육수요율, 분산 및 상대표준오차의 추정

$$\hat{R} = \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} x_{hij}}, \quad var_J(\hat{R}) = \frac{1}{k(k-1)} \sum_{\alpha=1}^k (\hat{R}_\alpha - \hat{R})^2,$$

$$\widehat{CV}(\hat{R}) = \frac{\sqrt{var_J(\hat{R})}}{\hat{R}} \times 100 (\%).$$

- 층별 아동의 보육수요율, 분산 및 상대표준오차의 추정

$$\widehat{R}_h = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} x_{hij}}, \quad \text{var}_J(\widehat{R}_h) = \frac{1}{k_h(k_h - 1)} \sum_{\alpha=1}^{k_h} (\widehat{R}_{\alpha h} - \widehat{R}_h)^2,$$

$$\widehat{CV}(\widehat{R}_h) = \frac{\sqrt{\text{var}_J(\widehat{R}_h)}}{\widehat{R}_h} \times 100(\%).$$

제5장 가중치조정

- 가중치 조정단계에서는 다음과 같은 사항을 고려함.
 - 기본가중치
 - 무응답조정가중치
 - 사후층화 가중치
 - 가중치 효과
 - 가중치의 절단
 - 가중치 조정시 기타 고려사항

제6장 조사 및 비표본오차 관리

- 조사관리
- 비표본오차 종류 및 관리
 - 포함오차
 - 무응답오차
 - 측정오차
 - 처리오차
 - 조정오차

제1장 서론

제1절 연구 목적 및 배경

1. 연구배경

『전국 보육실태조사』는 국가승인통계(승인번호: 제33107호)로서 2005년에 발효된 영유아 보육법에 따라 매 5년마다 실태조사를 통해 급격히 변화하고 있는 자녀의 양육환경에 부응하여 각종 영유아 관련 정책수립의 준거가 될 수 있는 객관성 있는 기초 자료의 생산에 목적을 두고 있다.

『전국보육실태조사』는 한국보건사회연구원에서 제한된 환경에서 2001년 전국표본 100개 지역 약 12,000가구에 거주하는 초등학교 이하 아동의 보육 및 교육서비스 이용실태를 파악하여 연령별, 지역별 보육수요율을 산출하여 중기 영유아 보육수요율을 34.1%로 산정하는 등의 연구결과를 제시하여 보건복지부의 보육정책을 위한 기초 자료로 활용한 바 있다.

2008년 6월 현재 전국의 보육시설은 총 32,149개소이며, 이중 국공립 보육시설은 1,769개소로 전체의 5.5%를 차지하며, 법인시설은 1,482개소로 4.6%, 민간시설은 14,127개소로서 43.9%를 차지하는 것으로 나타났다. 한편 가정보육시설은 14,375개소로 전체의 44.7%를 차지하여 민간보육시설이 대부분을 점유하고 있는 것으로 나타났다. 이를 지역별로 살펴보면, 서울과 경기지역이 각각 5,519개소와 9,233개소로 전체의 45.9%를 차지하고 있으며, 광역시 지역의 경우 울산을 제외한 5개 지역의 평균 보육시설 수는 1,377개소로 나타나고 있다. 한편 경기도를 제외한 도지역의 경우 강원, 충북 및 제주지역이 상대적으로 적은 규모이며, 나머지 지역은 충남이 1,289개소, 전북이 1,426개소, 전남이 1,024개소, 경북이 1,623개소, 경남이 2,369개소로 나타

나고 있다.

설립주체별로 지역별 보육시설 현황을 살펴보면, 국공립시설의 경우 서울과 경기 지역이 988개소로 전체의 55.8%를 차지하여 과반 이상이 서울 및 수도권에 집중되어 있는 것으로 파악되었고, 민간시설의 경우 또한 6,152개소로 전체의 43.5%를 차지하는 것으로 나타났다.

최근 급격한 출산 수준의 감소는 자녀양육지원 정책을 강화하는 배경이 되고 있는데, 2005년 1.08명, 2006년 1.13명, 2007년 1.26명으로 2005년 이후 점차적으로 개선되고는 있지만, 매우 낮은 수준으로 이와 같은 낮은 출산율의 원인으로는 여성의 사회참여확대, 높은 양육부담, 자녀양육에 대한 여성의 부담, 자녀에 대한 노후 보장 기대의 감소, 결혼가치관의 변화 이 매우 큰 원인으로 파악되고 있기 때문에 이에 대한 적절한 정책 수립이 필요한 실정이다.

이와 같이 육아보육의 부담이 갈수록 증가되는 상황에서 정부의 육아보육실태에 대한 개관적인 통계자료의 산출을 통해 육아보육 수요자의 욕구 및 만족도를 파악하여 육아보육정책의 올바른 방향성을 설정하고, 육아보육 지원 예산의 합리적 집행을 통한 국가 재정의 효율화에 기여할 수 있을 것으로 사료된다.

2. 연구목적

2005년에 발효된 영유아 보육법에 따라 매 5년마다 실태조사를 통해 급격히 변화하고 있는 자녀의 양육환경에 부응하여 각종 영유아 관련 정책수립의 준거가 될 수 있는 객관성 있는 기초 자료의 생산에 목적을 두고 있다.

이를 위해 수요자 지향적인 보육 서비스 공급을 위한 자료를 산출하여 지역, 아동연령, 소득수준, 母의 취업 등 가정 및 아동특성별 보육서비스 이용 및 수용자의 욕구를 파악하고, 수용자 특성, 요구 및 만족도를 파악함으로써 제도개선을 위한 정책자료를 생산하고자 한다.

또한 효율적인 예산지원 및 사업량을 산출하기위해 육아비용에 대한 부모부담 정도 파악, 지역별, 연령별 보육 수요율 파악 등 기초 자료를 생산하고자 한다.

한편 육아관련 각종제도와 연계추진을 위한 기초 자료를 생산하고, 보육 관련시설의 존재 및 육아휴직제 실시 등에 따른 중복되는 수요를 조절하여 예산집행 및 제도운용의 효율성 도모에 기여하고자 한다.

마지막으로 각종 육아정책수립을 위한 정부 공식통계 자료를 확보하고, 향후 보육 정책 수립 및 각종 육아정책을 위한 공식통계를 제공함으로써 정부 정책의 타당성과 일관성을 보장하기 위함이다.

한편 2009년 전국보육실태조사의 과학적인 표본설계를 통해서 전국의 영유아 보육실태 및 욕구를 정확히 파악하여 국가의 영유아 보육 정책을 적절히 수립할 수 있도록 하며, 또한 영유아 보육관련 연구에도 기여할 수 있기를 기대된다.

제2절 연구내용 및 방법

본 연구에서는 『전국 보육실태조사』의 목적을 충실히 달성하기 위해서 조사의 정확성과 신뢰성을 높이하고자 한다. 표본이 모집단을 잘 대표하고, 또한 변화하는 모집단의 특성이 표본에 반영되도록 설계하며, 과거 표본설계의 한계점이나 문제점 등을 검토한다. 또한, 2004년도에 실시된 전국 보육·교육 실태조사 자료를 이용해서 영유아에 대한 보육 수요율 추정값과 이에 대한 상대표준오차를 층별로 계산한다. 또한 본 조사와 관련이 깊은 항목에 대해 2005년 인구주택총조사 90%전수조사 자료를 분석해서 새로운 표본설계에 반영하는 한편 전국의 지역별, 설립주체별 보육시설 현황을 파악하여 보육시설 분포를 반영한 표본설계를 통해 보다 정밀한 표본설계가 이루어지도록 한다. 이를 위해서 다음과 같은 단계에 걸쳐서 세부적인 내용을 다룬다.

1. 과거 조사 자료에 대한 분석

2004년도에 조사된 전국보육·교육실태조사 원시자료(raw-data)를 이용해서 주요 문항별 특성을 분석한다. 즉, 주요 문항별 추정값과 표준오차, 그리고 변동계수(CV) 등을 계산하고, 표본배정에 사용할 변수를 선정한다. 또한 기존의 표본설계 전 과정

을 상세히 검토해서 현행 조사의 문제점, 개선내용 등을 종합적으로 파악하여 새로운 표본설계에 반영한다.

2. 모집단 분석

2009년도의 보육실태조사에 대한 표본설계에서는 목표모집단과 가장 근사한 2005년 인구주택총조사 90%전수조사 자료를 조사모집단으로 사용하였다.

인구주택총조사는 우리나라의 모든 가구와 가구원을 대상으로 한 5년 주기의 전수조사(complete survey, census)로 가구단위 조사에 가장 좋은 추출틀(sampling frame)이 된다. 특히 인구주택총조사를 위해 조사구를 설정하므로 본 조사의 경우도 조사구를 하나의 집락으로 사용한다. 조사구, 가구 그리고 인구 등에 대한 지역별, 동부, 읍·면부 분포와 조사구 형태별 분포 및 인구의 연령대 분포를 조사구별로 파악하여 표본배정에 활용하면 대표성 있는 표본을 얻을 수 있다. 또한 인구주택총조사 자료(또는 추계인구자료)는 추정 시 가중치 산정에 사용된다.

3. 2009년 표본설계

2009년 새로운 표본설계에서의 표본규모는 2004년도와 유사하게 전국 15,000가구(250개 조사구) 수준을 유지한다. 그러나 선행 조사의 경우와 마찬가지로 전국적인 보육수요율을 추정하여 보육정책에 반영코자 하는 것이 목적이므로, 가능한 지역적으로 영유아 출현율을 높이는 방향으로 설계가 이루어지도록 하며, 새로운 표본설계에서는 기존의 표본설계에서의 미흡했던 점이나 한계점들을 찾아서 보완토록 한다. 또한 2004년도의 보육실태조사자료와 2005년도 인구주택총조사의 90%전수조사 자료를 심층 분석해서 주요 관련변수를 찾고 이를 표본배정에 반영한다. 최종적으로 연구진, 외부전문가 그리고 보건복지가족부 실무진들과의 협의를 통해 표본배정 기준을 정한다.

추정에서는 가중치 계산과정을 다루고 추정치와 추정치의 오차공식을 유도한다. 또한 조사의 효율을 높이기 위해서 현행 조사 체계를 검토하여 문제점을 개선하는

방안을 마련한다.

표본추출은 동부와 읍면부를 기준으로 대도시, 중소도시, 농어촌의 3개 층으로 구성하고, 표본배정 원칙에 따라 표본추출프로그램(SAS)을 이용하여 배정한다. 표본조사구는 각 층별로 가구수(또는 가구원수)를 기준으로 확률비례계통추출 한다. 그리고 표본으로 선정된 조사구내에서 일정 수의 표본가구는 계통 추출한다.

4. 추정식 유도

2009년 표본설계에서는 2004년도 표본설계와 마찬가지로 3개 층별 추정과 전국 추정을 위한 추정치와 추정치의 오차계산을 위해 분산 추정식을 유도하고 아울러 가중값 계산과정을 다룬다. 그리고 층별 표본배정을 위한 배정공식을 제시한다.

5. 기타 사항

정확하고 효율적인 조사를 위해 모집단의 변동을 표본에 반영할 수 있도록 모집단과 표본의 변동에 따른 가중치 조정 문제를 중점적으로 다룬다. 특히 조사 후 추정 과정에서 널리 사용되고 있는 가중치 및 조정방법들을 상세히 다룬다.

상기의 단계적인 세부 내용을 수행하기 위해서 다음과 같은 연구방법을 사용한다.

모집단 특성 파악, 추출틀 확보, 조사항목에 대한 분석, 표본오차의 추정을 위하여 연관된 조사 자료를 수집한다. 또한 추정량, 추정량의 오차를 계산하기 위한 가중값 계산, 표본추출방법, 표본배정 공식, 분산 추정식 등에 대한 통계이론들을 살펴본다. 문헌검토 및 자료 수집을 통해서 여러 사례들을 검토한다.

본 조사를 총괄적으로 담당한 한국보건사회연구원 및 보건복지가족부의 담당자들과 보육관련 통계를 생산하고 있는 보건복지가족부, 한국보건사회연구원의 연구자들, 그리고 표본설계 전문가들과 자문회의 등을 통해 수시로 문제점 등을 협의한다. 과거 조사 자료의 분석, 새로운 모집단에 대한 특성 파악, 통계분석, 표본조사구 및 표본가구 선정, 추정치, 추정오차 계산을 위해서 EXCEL과 통계프로그램인 SAS, SPSS 등을 사용한다.

제3절 표본설계의 기본 목표와 절차

새로운 표본설계에서는 다음 사항들을 기본목표로 한다.

- 전국을 대도시, 중소도시, 농어촌으로 층화하여 영유아의 보육수요율을 정확히 파악할 수 있도록 표본 설계한다.
- 조사연구의 목적에 가장 부합되는 조사모집단을 설정하고 최선의 추출틀을 사용하여 표본이 모집단의 특성을 잘 대표할 수 있도록 한다.
- 과거 표본설계의 장·단점을 분석하여 문제점을 보완해서 조사결과의 정확성과 신뢰성을 높인다.
- 확률추출을 근간으로 해서 전국단위의 보육수요율 및 대도시, 중소도시, 농어촌별 추정이 가능하도록 추정량과 추정량의 분산 추정식을 유도한다.
- 조사 후 추정 시 문제가 되는 가중치 및 조정방법을 상세히 살펴본다. 그리고 과거 조사과정을 검토해서 조사의 효율성을 높이는 방안 등을 살펴본다.

한편, 효율적인 표본설계를 위해서 다음과 같은 단계적 절차를 밟는다.

- 과거 표본설계의 장·단점을 분석한다.
- 2004년 전국 보육·교육실태조사 자료로부터 대도시, 중소도시, 농어촌의 항목별 특성을 분석하고 표본배정을 위한 변수 등을 찾는다.
- 2005년 인구주택총조사의 90%전수조사 자료에 대한 상세한 분석을 하여, 본 조사의 목적에 부합되는 특성들과 전국 보육시설 현황을 검토하여 표본설계에 반영한다.
- 주어진 조사여건 하에서 층화방법, 추출방법, 표본배정방법 등을 결정한다.
- 가중치와 추정량, 그리고 추정량에 대한 오차공식 등을 유도한다.
- 표본조사구와 예비표본조사구를 추출하고, 표본조사구내에서 표본가구의 추출방법과 표본대체 방법 등을 제시한다.
- 다양한 가중치 및 조정방법을 다룬다. 그리고 향후 조사의 개선방안을 살펴본다.

제2장 2004년 조사에 대한 분석

제1절 2004년 조사의 개요

1. 조사 개요 및 방법

2004년 전국 보육·교육실태조사는 정부가 육아정책을 추진함에 있어 필요한 각종 정책수립의 준거가 될 수 있는 객관성 있는 자료 생산에 기본적인 목적이 있으며, 구체적인 목적은 다음과 같다.

첫째, 수요자 지향적인 보육 서비스 공급을 위한 자료를 산출하여 지역, 아동연령, 소득수준, 모의 취업 등 가정 및 아동특성별 보육서비스 이용 및 수용자의 욕구를 파악하고, 둘째, 수용자 특성, 요구 및 만족도를 파악함으로써 제도개선을 위한 정책 자료를 생산하고, 또한 효율적인 예산지원 및 사업량을 산출하기위해 육아비용에 대한 부모부담 정도 파악 및 지역별, 연령별 보육 수요율 파악과 더불어 육아관련 각종제도와 연계추진을 위한 기초 자료를 생산하고, 셋째, 보육 관련시설의 존재 및 육아휴직제 실시 등에 따른 중복되는 수요를 조절하여 예산집행 및 제도운용의 효율성 도모에 기여하고, 넷째, 각종 육아정책수립을 위한 정부 공식통계 자료를 확보하고, 향후 보육정책 수립 및 각종 육아정책을 위한 공식통계를 제공함으로써 정부 정책의 타당성과 일관성을 보장하기 위함이다.

2004년 조사의 경우 조사원에 의한 면접조사로 진행되었으며, 표본으로 추출된 가구를 대상으로 가구조사표를 이용하여 가구원 및 가구 특성에 대한 조사를 실시한 후 초등학교 이하 아동이 있는 경우 해당 아동에 대해 개별아동 조사를 실시하였다. 조사내용은 여성부, 교육인적자원부, 대통령자문 고령화 및 미래사회위원회 정책업무 담당자 및 대학교수, 보육시설 장 등으로 구성된 자문단과의 수차례에 걸친 회의

를 통해 검토하였다.

2. 조사표의 구성

조사내용을 주요 영역별로 살펴보면 다음과 같다.

가. 가구조사

□ 가구의원 특성

- 가구의원의 성명, 가구주와의 관계, 성별, 연령, 생년월일, 교육정도 및 졸업여부, 결혼상태, 취업상태, 건강상태 등.
- 가구의 주택 및 경제상태
- 가구원 및 보육·교육대상 아동 수, 주택유형 및 소유형태, 주 수입원, 월평균 가구소득 및 지출 등.
- 초등학교 이하 아동가구 대상 부모 특성 및 정책관련 의견
- 초등학교 이하 아동 부모의 종사상 지위, 근로시간, 월수입 등에 대한 특성
- 정부 지원 및 부모전담 수준 등 보육·교육정책에 대한 의견
- 보육·교육 대상 아동에 관한 사항
- 초등학교 이하 아동의 현재 기관, 개인 등 20개의 보육 및 교육서비스 이용현황

나. 개별 아동조사

- 보육·교육 이용경험
- 주 양육자 : 아동 연령대별 주 양육자 및 필요 양육자
- 보육·교육 서비스 이용경험 : 최초 서비스 제공기관(자), 이용이유
- 보육·교육 기관이용현황
- 기관입지 : 소재지, 운영주체, 주변지역 현황

- 기관이용형태 : 이용일수, 이용시간, 희망운영시간, 이용이유, 이용시기, 기관선택사고려사항, 취약보육이용여부 및 비용, 등하원 방법 및 시간
- 서비스제공 : 기관에서 강조하는 내용, 부모가 중요시하는 내용, 부모서비스 제공여부, 빈도, 특기교육
- 보육료 : 보육료, 부담정도, 적정보육료 수준 등
- 만족도 : 이용만족도, 불만족 사유
- 의견사항 : 계속이용의사 등 보육시설운영위원회 도입에 관한 의견 등
- 일반학원 이용
- 이용현황 : 이용하는 학원수, 주평균이용시간, 월평균 비용 등
- 인식 : 시설설비, 강사, 교육내용 수준 및 비용에 관한 인식, 만족도
- 선택기준 : 학원선택시 중요시하는 내용
- 개별지도, 그룹지도 및 학습지 활용교육
- 이용 회수, 이용시간, 비용, 만족도
- 개인양육지원 서비스 이용
- 보육서비스 제공자, 보육장소, 보육서비스 제공빈도, 월평균 시간, 보육비용, 애로사항, 보육시설 미 이용 이유, 만족도, 비공식보육 계속의사 등
- 초등학생 방과 후 보육·교육
- 방과 후 보육실태 및 관련 의견
- 초등학교 방과 후 특기적성 및 보육 프로그램 이용 및 욕구
- 영유아 보육·교육 욕구 등
- 보육시설 이용의사 : 보육시설 미이용 사유 및 향후 이용의사
- 유치원 이용의사 : 유치원 미이용 사유 및 향후 이용의사
- 시간 연장형 등 취약 보육 욕구

제2절 2004년 조사의 특징

2004년 1월 개정된 영유아 보육법에 따라 정부 및 지방정부의 보육수요를 파악하여 보육계획을 수립하도록 정하고 있으며, 매 5년마다 전국보육실태조사를 실시하도록 하고 있다.

2004년 전국보육·교육실태 조사에서는 전국 200개 조사구 약 12,000가구에 대해 가구원 및 가구의 특성을 파악하여 초등학교 이하 아동이 있는 경우 이들에 대한 개개인의 개별 조사를 실시하였다. 한편 조사완료율 및 아동 출현율이 저조하여 2004.10.20~12.10 까지 20개 조사구를 추가로 조사하여 최종 조사를 완료하였다.

2004년에는 인구주택총조사 10% 조사구 자료를 이용하여 전국을 3개의 층으로 층화하여 각 층에서 주택 특성에 따라 정렬한 후 200개 조사구를 표본으로 추출하였다. 조사대상 아동은 지역 조사구 대상 가구 내에 실제로 거주하는 영유아 및 초등학교 이하 가구원으로 정의하고 있다.

한편 조사목적에 따라 영유아 및 조사대상 아동을 다음과 같이 정의하고 있다.

- 영아 : 출생 후 36개월 미만 아동
- 유아 : 36개월 이상 만 6세 미만의 아동을 의미함. 단, 2004년 조사에서는 36개월 이상 취학 전 아동을 모두 의미하는 용어로 사용하였음.
- 영유아 : 출생 후 만 6세 미만의 아동을 의미함. 단, 2004년 조사에서는 미취학아동과 같은 대상을 지칭하는 의미로 사용됨.
- 미취학아동 : 초등학교 취학 전 아동 전체를 의미하며, 3월1일 기준으로 만 6세 아동은 취학 시점이지만, 부모의 취학유예 의사에 따라 취학을 유예할 수 있기 때문에 만 6세 이상 아동도 포함됨.
- 취학 전 아동 : 취학 전 1년인 만 5세 이상의 아동으로 초등학교에 입학하지 않은 아동을 의미하며, 만 6세 이상 아동도 포함이 가능함.

한편 2004년 조사의 경우 2002년과는 달리 보육 및 교육에 대한 실태조사를 수행하였고, 한국보건사회연구원을 포함한 4개 국책기관이 참여하여 보육시설과 유치원에 대한 실태조사가 별도로 수행되었으며, 표준 보육·교육 단가산출 연구가 수행되었다.

2004년 조사의 경우 200개 조사구에서 파악된 총 가구수는 11,960가구로 파악되었고, 아동은 6,692명으로 조사되었다. 이중 가구는 87.1%인 10,423가구, 아동은 95.8%인 6,412명이 조사 완료되었다.

제3절 2004년 표본설계 및 특성별 분석

1. 표본설계의 개요

전국을 모집단으로 하여 2000년 인구주택총조사의 10% 조사구인 25,300여개 조사구를 추출틀로 구성하여 서울 및 광역시(부산, 대구, 인천, 광주, 대전, 울산)의 동부, 기타시도의 동부, 기타시도의 읍면부 등으로 3개 층으로 구성하였다. 각 층에 대해서 조사구 특성별로 조사구내 주택 특성에 따라 ① 단독주택, ② 아파트, ③ 연립 및 다세대 주택, ④ 기타 주택 순으로 정렬하였다. 이와 같이 정렬된 조사구 이외에 2000년 인구주택 총조사 이후 아파트 신축 지역을 추가하여 구성된 전체 조사구에서 지역의 변동 상황을 고려하여 최종적으로 200개 조사구를 표본지역으로 추출하였다.

〈표 2-1〉 2004년 조사의 표본 조사구 특성

(단위: 개)

구 분	조사구 비율	표본조사구 추출결과	
		가구수	평균 가구수
서울의 동부	0.23	2,623	57
광역시의 동부	0.24	2,843	59
기타시의 동부	0.32	3,734	59
읍부	0.09	1,220	64

구 분	조사구 비율	표본조사구 추출결과	
		가구수	평균 가구수
면부	0.14	1,401	58
계	1.00	11,821	59

2. 표본설계의 특징

2004년 표본설계의 특징을 살펴보면 다음과 같다.

첫째, 표본 추출틀로 2000년 인구주택 총조사의 10% 조사구 자료를 이용하였다. 이는 조사구당 평균 가구수가 약 60가구로서 영유아 가구 출현율을 높이고, 90%자료에 비해 조사구의 정보를 보다 자세히 이용할 수 있다는 장점이 있다.

둘째, 2004년 표본설계에서 고려한 표본 배분 방식이나, 표본추출방식에 관한 사항이 언급되지 않았으나, 통상적으로 조사구 특성별로 정렬한 후 확률비례 계통 추출을 사용하였고, 별도의 표본배분 방식을 이용하기 보다는 3개 층에 대한 조사구 비율에 따라 적정한 추출률을 적용하여 최종적으로 표본조사구를 추출한 것으로 판단된다.

셋째, 16개 광역시도별 층화보다는 표본수의 한계 때문에 서울 및 광역시의 동부, 기타 시도의 동부, 기타시도의 읍면부 등 3개 층으로 층화하여 정도를 높인 것으로 사료된다.

넷째, 별도의 표본 규모에 대한 산정을 하지 않고, 초기에 정해진 200개 조사구를 3개 층에 적절히 배분한 것으로 판단된다.

다섯째, 2000년 인구주택총조사 10%자료를 이용하여 표본추출이 되었기 때문에 추출틀의 노후화에 따른 별도의 부가적인 추출틀을 활용하여 표본을 추출하였다. 이는 추출틀의 포괄성을 높인다는 측면에서는 적절한 절차로 판단되나, 그 규모가 전체 조사구에 비해 상대적으로 작기 때문에 추후 가중치 조정을 통해 보완할 수 있을 것으로 판단된다.

3. 2004년 조사결과의 주요 특성별 분석

2004년 조사 결과를 지역별로 살펴보면 다음과 같다.

〈표 2-2〉 2004년 조사결과의 지역별 특성

(단위: 개, 명)

지 역	가 구				아 동			
	전체	대도시	중소도시	읍면	전체	대도시	중소도시	읍면
서 울	1,381	1,381	-	-	647	647	-	-
부 산	807	762	-	45	438	416	-	-
대 구	473	422	-	51	301	244	-	-
인 천	510	510	-	-	347	347	-	-
광 주	368	368	-	-	305	305	-	-
대 전	297	397	-	-	178	178	-	-
울 산	351	307	-	44	243	226	-	17
경 기	1,998	-	1,549	449	1,270	-	1,019	251
강 원	497	-	244	253	254	-	122	132
충 북	499	-	244	253	256	-	130	126
충 남	588	-	209	379	408	-	159	149
전 북	537	-	305	232	382	-	225	157
전 남	603	-	336	167	390	-	203	187
경 북	573	-	308	165	366	-	209	157
경 남	763	-	426	337	499	-	314	185
제 주	178	-	124	54	128	-	88	40
계	10,423	4,047	3,745	2,631	6,412	2,363	2,469	1,580

다음으로 16개 광역시도별로 조사대상 아동수에 대한 지역별 변동계수를 산출하면 다음과 같다.

표본수가 가장 적은 제주도의 경우 약 10%로 가장 높고, 다음은 대전지역으로 8.6% 정도로 나타나고 있으며, 경기도가 3.3%정도로 가장 낮게 나타나고 있다(표 2-3 참조). 따라서 16개 광역시도별 아동수 또는 아동 출현율에 대한 통계산출은 지

역별로 CV 값의 변동이 크기 때문에 곤란할 것으로 사료되며, 결과적으로 지역을 5개 권역 또는 3개 권역으로 통합하여 산출하는 것이 바람직할 것으로 판단된다. 또한 표본설계단계에서 16개 광역시도별 통계산출을 위한 층화를 고려하지 않았기 때문에 이와 같은 결과는 필연적이라 할 수 있다.

이와는 별도로 5개 권역별 아동수에 대한 CV와 3개 권역별 CV를 각각 산출한 결과 16개 광역시도별 CV보다는 매우 안정적으로 나타나고 있으며, 결과적으로 5개 권역 또는 3개 권역별 통계를 산출할 경우 전반적인 통계값들이 안정적으로 나타날 것으로 기대된다(표 2-4, 표 2-5 참조).

〈표 2-3〉 16개 광역시도별 조사대상 아동 수에 대한 지역별 변동계수

(단위: %)

지 역	변동계수(CV)	지 역	변동계수(CV)
서 울	4.5226	강 원	8.5302
부 산	5.6709	충 북	8.262
대 구	6.9637	충 남	6.8249
인 천	5.8499	전 북	6.7735
광 주	6.661	전 남	7.0282
대 전	8.6113	경 북	6.8127
울 산	7.2591	경 남	5.5771
경 기	3.3447	제 주	9.9175

〈표 2-4〉 5개 권역별 조사대상 아동 수에 대한 변동계수

(단위: %)

층 별	변동계수(CV)
서울시의 동	1.8618
광역시의 동	1.1968
기타시의 동	1.0225
읍	1.9564
면	2.9760

〈표 2-5〉 3개 권역별 조사대상 아동 수에 대한 변동계수

(단위: %)

층 별	변동계수(CV)
대도시	2.3795
중소도시	2.3098
읍면	3.8020

다음으로 조사대상 가구의 주택소유 형태별로 층화 기준에 따라 0~6세 아동출현율과 그에 따른 변동계수를 추정하였다. <표 2-6>으로부터 각각의 층화 기준에 따른 전국의 0~6세 아동출현율에 대한 추정치는 동일하지만, 각 기준별 변동계수는 약간씩 차이가 나고 있다.

〈표 2-6〉 층화기준에 따른 주택소유형태별 아동출현율 추정치와 변동계수

(단위: %)

층화 기준	주택소유형태	아동출현율	표준오차	변동계수(CV)
16개 지역 (16개 특광역시)	자가	23.86	0.72	3.03
	전세	39.81	1.57	3.94
	전월세	25.33	1.62	6.38
	사글세	17.64	3.30	18.70
	무상	28.40	2.90	10.22
5개 권역 (서울, 광역시, 기타시의 동, 읍, 면)	자가	23.86	0.72	3.02
	전세	39.81	1.57	3.94
	전월세	25.33	1.64	6.49
	사글세	17.64	3.33	18.88
	무상	28.40	2.90	10.19
3개 권역 (대도시, 중소도시, 읍면)	자가	23.86	0.72	3.03
	전세	39.81	1.57	3.94
	전월세	25.33	1.65	6.52
	사글세	17.64	3.33	18.87
	무상	28.40	2.89	10.18

변동계수를 기준으로 볼 때, 전국추정치의 안정성은 5개 권역별 추정치가 가장 안정적으로 나타난 반면 3개 권역별 추정치가 가장 높게 나타나고 있으나, 그 차이는 매우 미미하다. 결과적으로 층화기준을 어떻게 설정하는지 간에 주택 소유형태별로 아동수 추정에는 별 영향을 끼치지 않는 것으로 볼 수 있다.

다음은 16개 광역시도별로 주택소유형태별 아동출현율을 추정한 결과 <표 2-7>과 같이 서울 자가의 경우 아동출현율은 19.73%, 전세는 32.68%, 전월세는 13.09%, 사글세는 12.29%, 무상은 20.62%로 나타나고 있으며, 각각의 출현율에 따른 변동계수는 자가에서는 9.46%, 전세는 9.21%, 전월세는 22.18%, 사글세는 68.98%, 무상은 42.73%로 추정되고 있다. 지역적으로 볼 때 특·광역시 지역의 자가에 대한 아동출현율에 대한 변동계수는 10.71~16.42%로 분포의 폭이 넓은 것으로 나타나고 있으며, 대전의 경우 16.42%로 가장 높게 나타나고 있다. 반면에 다른 소유형태의 경우 전반적으로 변동이 매우 심한 것을 알 수 있다. 한편 전세의 경우 아동 출현율의 변동계수를 살펴보면 서울은 9.21%, 부산은 13.66%, 대구는 17.52% 등으로 나타나고 있다.

<표 2-7> 16개 광역시도별 주택소유 비율에 따른 아동출현율에 대한 변동계수
(단위: %)

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
서울	자가	19.37	1.83	9.46
	전세	32.68	3.01	9.21
	전월세	13.09	2.90	22.18
	사글세	12.29	8.48	68.98
	무상	20.62	8.81	42.73
부산	자가	24.80	2.66	10.71
	전세	39.64	5.42	13.66
	전월세	14.94	4.27	28.58
	사글세	49.59	26.16	52.75
	무상	39.03	22.60	57.92

〈표 2-7〉 계속

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
대구	자가	26.28	3.27	12.43
	전세	36.57	6.41	17.52
	전월세	18.50	6.63	35.84
	사글세	28.40	21.68	76.35
	무상	7.37	7.60	103.07
인천	자가	26.68	3.49	13.10
	전세	41.69	8.08	19.38
	전월세	43.09	7.07	16.41
	사글세	0.00	0.00	-
	무상	19.33	13.19	68.27
광주	자가	35.37	4.33	12.23
	전세	62.29	8.74	14.03
	전월세	0.00	0.00	-
	사글세	25.00	14.35	57.41
	무상	41.57	13.59	32.69
대전	자가	24.95	4.10	16.42
	전세	44.41	8.31	18.72
	전월세	13.58	6.92	50.93
	사글세	0.00	0.00	-
	무상	0.00	0.00	-
울산	자가	26.26	3.76	14.31
	전세	39.46	10.32	26.16
	전월세	19.13	6.56	34.26
	사글세	0.00	0.00	-
	무상	45.77	24.33	53.16
경기	자가	27.35	1.82	6.66
	전세	39.85	3.18	7.97
	전월세	28.56	4.15	14.52
	사글세	8.32	4.84	58.08
	무상	40.40	9.66	23.92

〈표 2-7〉 계속

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
강원	자가	14.30	2.32	16.23
	전세	36.53	11.30	30.94
	전월세	39.74	12.71	31.99
	사글세	0.00	0.00	-
	무상	34.75	8.23	23.67
충북	자가	16.11	2.91	18.04
	전세	37.64	6.83	18.15
	전월세	31.19	6.06	19.42
	사글세	21.06	15.14	71.91
	무상	20.39	9.79	48.02
충남	자가	20.27	2.78	13.71
	전세	54.78	8.65	15.79
	전월세	60.96	10.18	16.70
	사글세	14.56	8.23	56.53
	무상	32.35	10.13	31.31
전북	자가	32.27	3.97	12.29
	전세	40.37	7.04	17.44
	전월세	20.85	6.34	30.43
	사글세	46.44	19.28	41.51
	무상	24.60	11.69	47.52
전남	자가	15.77	2.91	18.47
	전세	46.46	8.99	19.34
	전월세	31.44	9.57	30.45
	사글세	28.64	13.64	47.64
	무상	15.49	7.68	49.61
경북	자가	28.00	3.32	11.87
	전세	35.34	7.90	22.36
	전월세	31.57	11.27	35.71
	사글세	8.25	5.92	71.73
	무상	30.93	8.19	26.46

〈표 2-7〉 계속

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
경남	자가	20.15	2.06	10.24
	전세	56.01	10.03	17.90
	전월세	40.27	12.35	30.66
	사글세	11.99	9.77	81.48
	무상	21.53	9.04	41.99
제주	자가	27.40	5.12	18.68
	전세	43.81	19.08	43.54
	전월세	137.07	42.04	30.67
	사글세	0.00	0.00	-
	무상	15.57	11.26	72.33

다음으로 전국을 5개 권역으로 층화하여 주택소유형태별 아동 출현율에 대한 변동계수를 추정해본 결과 전체적으로 자가 비율에 대한 변동계수가 4.55~14.90%로 변동이 심하며, 전세의 경우 아동출현율은 읍 면 지역에서 각각 11.11%와 17.67%로 타지역에 비해 높게 나타났다. 전월세의 경우 서울은 22.18%, 광역시는 13.20%, 기타시에서는 9.67%로 나타난 반면, 읍지역과 면지역은 각각 14.43%와 30.26%로 상대적으로 높게 나타났다. 그 외 주택소유형태 따른 아동출현율은 지역별로 매우 큰 변동을 나타내고 있음을 알 수 있다(표 2-8 참조).

〈표 2-8〉 5개 권역별 주택소유형태에 따른 아동출현율에 대한 변동계수

(단위: %)

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
서울의 동	자가	19.37	1.83	9.46
	전세	32.68	3.01	9.21
	전월세	13.09	2.90	22.18
	사글세	12.29	8.48	68.98
	무상	20.62	8.81	42.73

〈표 2-8〉 계속

(단위: %)

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
광역시·시의 동	자가	26.94	1.46	5.40
	전세	42.84	3.08	7.18
	전월세	21.58	2.85	13.20
	사글세	27.17	9.52	35.05
	무상	27.25	7.64	28.03
기타시의 동	자가	26.91	1.22	4.55
	전세	40.36	2.57	6.36
	전월세	32.62	3.15	9.67
	사글세	17.41	5.41	31.10
	무상	37.19	4.87	13.09
읍	자가	26.22	2.08	7.92
	전세	45.39	5.04	11.11
	전월세	47.32	6.83	14.43
	사글세	14.36	5.82	40.55
	무상	25.07	6.48	25.86
면	자가	12.38	1.85	14.90
	전세	46.78	8.26	17.67
	전월세	18.22	5.51	30.26
	사글세	12.00	7.09	59.08
	무상	21.49	6.39	29.75

한편 <표 2-9>는 전국을 3개 권역으로 층화하여 주택소유형태별 아동출현율과 그에 따른 변동계수를 추정한 결과이다. <표 2-7>과 <표 2-8>의 결과에 비해 각 층별로 변동계수 값이 매우 안정적으로 나타나고 있다. 각 층별로 자가에 대한 아동출현율의 변동계수가 각각 4.70%, 4.55%, 7.44%로 안정적이며, 전세의 경우 5.68%, 6.36%, 9.67%로 나타나 자가에 비해서는 약간의 변동이 있지만, 전반적으로는 안정적이라 할 수 있다.

〈표 2-9〉 3개 권역별 주택소유형태에 따른 아동출현에 대한 변동계수

(단위: %)

지 역	주택소유형태	아동출현율	표준오차	변동계수(CV)
대도시	자가	24.46	1.15	4.70
	전세	38.16	2.17	5.68
	전월세	18.17	2.07	11.39
	사글세	23.10	7.31	31.67
	무상	24.81	5.80	23.37
중소도시	자가	26.91	1.22	4.55
	전세	40.36	2.57	6.36
	전월세	32.62	3.15	9.67
	사글세	17.41	5.41	31.10
	무상	37.19	4.87	13.09
읍면	자가	18.57	1.38	7.44
	전세	45.93	4.44	9.67
	전월세	36.11	4.84	13.41
	사글세	13.46	4.48	33.27
	무상	23.08	4.56	19.75

다음으로 <표 2-10>에서 주택 유형별 아동출현율에 대한 전국 추정치를 층화기준에 따라 분석해보면 층화별 주택유형별 아동출현율은 동일하게 추정되지만, 각각의 층화방법에 따라 변동계수 값이 약간씩 차이가 남을 알 수 있다. 전국 추정에서 가장 안정적으로 추정되는 층화 방법은 5개 권역별 층화와 3개 권역별 층화로 나타났으며, 제한된 표본수하에서 지역별 세분화에 따른 각 지역별 통계의 안정성에 문제가 되기 때문에 16개 특·광역시별 층화는 문제가 발생할 수 있다.

〈표 2-10〉 층화기준에 따른 주택유형별 아동출현율에 추정치와 변동계수
(단위: %)

층화 기준	주택유형	아동출현율	표준오차	변동계수(CV)
16개 지역 (16개 특광역시)	단독주택	18.24	0.92	5.04
	아파트	37.04	0.91	2.46
	연립주택	24.43	2.70	11.03
	다세대	30.38	2.10	6.93
	상가주택	24.59	3.64	14.81
5개 권역 (서울, 광역시, 기타시의 동, 읍, 면)	단독주택	18.24	0.92	5.04
	아파트	37.04	0.92	2.47
	연립주택	24.43	2.69	11.02
	다세대	30.38	2.11	6.94
	상가주택	24.59	3.67	14.93
3개 권역 (대도시, 중소도시, 읍면)	단독주택	18.24	0.92	5.04
	아파트	37.04	0.92	2.47
	연립주택	24.43	2.69	11.01
	다세대	30.38	2.11	6.94
	상가주택	24.59	3.69	14.99

다음의 분석에서는 주택유형별 아동출현율과 그에 따른 변동계수를 세 가지 층화기준에 따라 추정한 결과를 보여주고 있다.

우선 <표 2-11>에서는 16개 특·광역시별 주택 유형에 따른 아동출현율과 그에 따른 변동계수를 살펴보면 지역별로 매우 불안정한 값을 나타내고 있으며, 광주, 대전, 울산 등 특정지역의 경우 연립, 다세대 주택유형이 파악되지 않아 추정이 불가능한 경우도 발생하고 있다. 따라서 주어진 표본수하에서 16개 지역별 통계산출은 통계적으로 유의미하지 않은 결과를 산출할 우려가 있는 것으로 나타났다.

전반적으로 특·광역시 지역에서는 단독주택보다는 아파트에서의 아동출현율의 변동이 작게 나타나며, 이와 같은 현상은 다른 지역에서도 동일한 결과를 나타내고 있다.

〈표 2-11〉 16개 광역시도별 주택유형에 따른 아동출현율에 대한 변동계수
(단위: %)

지 역	주택유형	아동출현율	표준오차	변동계수(CV)
서울	단독주택	14.48	2.00	13.78
	아파트	29.42	2.35	8.00
	연립주택	17.48	5.09	29.12
	다세대	29.03	4.53	15.61
	상가주택	39.89	10.48	26.26
부산	단독주택	23.00	4.09	17.80
	아파트	31.77	3.03	9.53
	연립주택	17.27	7.15	41.41
	다세대	31.35	6.33	20.20
	상가주택	4.50	4.50	99.86
대구	단독주택	20.70	4.16	20.10
	아파트	39.77	4.51	11.34
	연립주택	0.00	0.00	0.00
	다세대	22.78	7.10	31.19
	상가주택	24.67	10.93	44.33
인천	단독주택	25.78	6.84	26.53
	아파트	34.97	3.63	10.38
	연립주택	30.56	9.61	31.45
	다세대	43.72	11.65	26.65
	상가주택	14.24	10.40	73.03
광주	단독주택	19.08	5.50	28.80
	아파트	49.25	4.51	9.15
	연립주택	0.00	0.00	0.00
	다세대	0.00	0.00	0.00
	상가주택	66.67	33.33	50.00
대전	단독주택	15.67	5.00	31.88
	아파트	34.05	4.24	12.44
	연립주택	0.00	0.00	0.00
	다세대	0.00	0.00	0.00
	상가주택	0.00	0.00	0.00

〈표 2-11〉 계속

지 역	주택유형	아동출현율	표준오차	변동계수(CV)
울산	단독주택	20.63	4.57	22.17
	아파트	35.87	4.66	12.99
	연립주택	0.00	0.00	0.00
	다세대	28.57	19.41	67.94
	상가주택	30.05	21.37	71.12
경기	단독주택	24.04	2.57	10.68
	아파트	37.45	2.11	5.63
	연립주택	32.08	5.01	15.61
	다세대	27.52	3.71	13.48
	상가주택	41.10	15.81	38.47
강원	단독주택	11.27	3.64	32.24
	아파트	31.80	3.78	11.89
	연립주택	15.43	5.88	38.09
	다세대	0.00	0.00	0.00
	상가주택	7.62	7.68	100.79
충북	단독주택	14.19	3.05	21.50
	아파트	43.86	4.65	10.61
	연립주택	0.00	0.00	0.00
	다세대	8.18	8.26	101.02
	상가주택	5.18	5.19	100.10
충남	단독주택	14.67	3.01	20.53
	아파트	68.68	5.44	7.92
	연립주택	29.79	10.05	33.73
	다세대	21.74	10.81	49.72
	상가주택	37.85	18.37	48.54
전북	단독주택	21.31	4.83	22.65
	아파트	42.38	3.91	9.22
	연립주택	0.00	-	-
	다세대	35.98	13.99	38.87
	상가주택	10.88	10.78	99.05

〈표 2-11〉 계속

지 역	주택유형	아동출현율	표준오차	변동계수(CV)
전남	단독주택	13.34	3.93	29.47
	아파트	30.77	3.46	11.25
	연립주택	0.00	0.00	0.00
	다세대	0.00	0.00	0.00
	상가주택	28.92	28.86	99.80
경북	단독주택	19.40	3.32	17.11
	아파트	49.97	4.87	9.75
	연립주택	0.00	0.00	0.00
	다세대	25.53	8.83	34.59
	상가주택	36.01	20.71	57.50
경남	단독주택	15.48	3.18	20.55
	아파트	33.97	3.22	9.47
	연립주택	0.00	0.00	-
	다세대	28.42	6.65	23.38
	상가주택	19.49	10.41	53.42
제주	단독주택	5.26	3.67	69.75
	아파트	25.00	6.25	24.99
	연립주택	0.00	0.00	0.00
	다세대	54.85	10.31	18.80
	상가주택	0.00	-	-

<표 2-12>에는 전국을 5개 권역으로 총화하여 주택유형별 아동출현율과 변동계수를 나타내고 있다. 전체적으로 16개 지역별 총화 때 보다 주택유형에 따른 아동출현율이 상대적으로 안정된 결과를 나타내고 있으며, 특히 면지역을 제외하고 아파트에서 아동출현율에 대한 변동계수 값은 매우 안정된 결과를 보이고 있다.

〈표 2-12〉 5개 권역별 주택유형에 따른 아동출현율에 대한 변동계수

지 역	주택유형	아동출현율	표준오차	변동계수(CV)
서울의 동	단독주택	14.48	2.00	13.78
	아파트	29.42	2.35	8.00
	연립주택	17.48	5.09	29.12
	다세대	29.03	4.53	15.61
	상가주택	39.89	10.48	26.26
광역시의 동	단독주택	21.54	2.09	9.69
	아파트	36.50	1.64	4.48
	연립주택	24.73	6.25	25.28
	다세대	32.21	4.94	15.32
	상가주택	20.28	5.83	28.75
기타시의 동	단독주택	22.69	1.90	8.36
	아파트	37.41	1.42	3.78
	연립주택	20.84	3.83	18.37
	다세대	28.58	3.00	10.49
	상가주택	25.12	7.39	29.43
읍	단독주택	16.70	2.45	14.68
	아파트	50.81	2.36	4.64
	연립주택	32.05	8.10	25.27
	다세대	37.62	7.58	20.15
	상가주택	19.31	11.68	60.48
면	단독주택	13.54	1.81	13.35
	아파트	45.73	7.54	16.48
	연립주택	46.67	13.33	28.57
	다세대	30.02	8.34	27.79
	상가주택	17.84	8.13	45.59

한편 다음의 <표 2-13>에는 전국을 3개 권역으로 층화하여 주택유형별 비율과 변동계수를 나타내고 있다. 전체적으로 16개 지역별 층화나 5개 권역별 층화에 따른 주택유형별 비율보다는 상대적으로 매우 안정된 결과를 나타내고 있으며, 특히 중소도시와 읍면지역의 상가주택 비율에 대한 변동계수가 29~36%이상으로 나타나고 있지만,

나머지 지역의 주택유형별 비율에 대한 변동계수는 안정된 결과를 보이고 있다.

〈표 2-13〉 3개 권역별 주택유형에 따른 아동출현율에 대한 변동계수 (단위: %)

지 역	주택유형	아동출현율	표준오차	변동계수(CV)
대도시	단독주택	18.70	1.49	7.95
	아파트	34.46	1.35	3.92
	연립주택	20.47	3.95	19.27
	다세대	30.56	3.34	10.92
	상가주택	27.08	5.32	19.63
중소도시	단독주택	22.69	1.90	8.36
	아파트	37.41	1.42	3.78
	연립주택	20.84	3.83	18.37
	다세대	28.58	3.00	10.49
	상가주택	25.12	7.39	29.43
읍면	단독주택	14.68	1.46	9.91
	아파트	50.38	2.25	4.46
	연립주택	36.38	6.93	19.05
	다세대	34.28	5.60	16.33
	상가주택	18.32	6.62	36.11

결과적으로 2004년 보육·교육실태조사의 경우 지역별 표본수가 작은 관계로 통계의 불안정성을 극복하고자 5개 권역별 층화 후 3개 권역에 대한 통계를 산출하여 하였으며, 이는 매우 적절한 방안으로 사료된다. 따라서 2009년의 표본설계 역시 2004년의 표본수와 거의 동일하기 때문에 16개 특·광역시별 층화 보다는 5개 권역이나 3개 권역별 층화 후 그에 따른 통계를 산출하는 것이 바람직 할 것으로 판단 된다.

한편 지역별 0~6세의 아동분포를 살펴보면 <표 2-14>와 같이 지역별 아동 연령별 분포를 살펴보면 경기지역이 19.95%로 가장 많은 수의 아동이 조사되었고, 제주지역 이 1.73%로 가장 적게 조사되었다.

〈표 2-14〉 지역별 아동 연령대별 분포

(단위: %)

지 역	계	0세	1세	2세	3세	4세	5세	6세
전 국	100.00	11.72	12.04	13.49	14.34	15.98	15.51	16.92
서 울	10.37	12.10	10.99	9.58	8.57	10.65	9.55	11.36
부 산	7.09	5.91	9.16	7.01	6.81	8.68	5.28	6.89
대 구	5.11	5.38	7.07	4.21	5.27	5.72	4.67	3.91
인 천	5.17	5.65	4.71	5.14	6.59	3.94	4.67	5.59
광 주	4.82	4.30	5.50	4.67	5.71	3.16	6.30	4.28
대 전	2.71	2.96	2.88	3.27	2.86	1.78	2.64	2.79
울 산	3.21	2.15	2.09	3.50	2.42	4.54	4.07	3.17
경 기	19.95	18.01	18.85	19.16	20.88	18.93	21.14	21.79
강 원	3.78	2.96	3.40	4.44	3.96	4.14	3.46	3.91
충 북	5.11	6.99	5.76	8.18	3.96	5.92	3.25	2.79
충 남	7.44	8.87	6.54	6.31	6.81	8.68	7.52	7.26
전 북	6.21	3.76	4.45	7.24	4.84	8.09	7.11	6.89
전 남	4.70	5.11	5.76	4.67	5.49	3.16	4.27	4.84
경 북	5.86	6.99	6.81	6.07	5.71	6.31	4.88	4.84
경 남	6.74	6.45	4.71	5.61	7.91	4.34	9.15	8.38
제 주	1.73	2.42	1.31	0.93	2.20	1.97	2.03	1.30

이와 더불어 전국을 5개 권역과 3개 권역으로 구분하여 0~6세 까지 아동 수의 분포를 살펴보면 <표 2-15>와 <표 2-16>과 같다. 5개권역의 경우 읍지역에 비해 면지역의 아동 비율이 낮게 나타나고 있으며, 서울과 대도시의 동지역의 경우 서울지역의 아동비율이 낮게 조사되었다. 3개권역의 경우 연령대별 아동비율이 전체적으로 고르게 나타나고 있으며, 읍면지역이 타지역에 비해 연령대별로 도 낮게 나타나고 있어 농어촌 지역의 아동 출현율이 상대적으로 도시지역에 비해 낮음을 알 수 있다.

〈표 2-15〉 5개 권역별 아동 연령별 분포

(단위: %)

5개 권역	계	0세	1세	2세	3세	4세	5세	6세
계	100.00	11.72	12.04	13.49	14.34	15.98	15.51	16.92
서울의 동	10.37	12.10	10.99	9.58	8.57	10.65	9.55	11.36
광역시의 동	26.19	24.46	29.32	26.64	27.47	24.85	25.61	25.51
기타시의 동	37.16	38.17	33.51	39.25	37.14	35.90	39.02	36.87
읍	20.64	22.58	19.37	19.86	20.66	22.49	20.12	19.55
면	5.64	2.69	6.81	4.67	6.15	6.11	5.69	6.70

〈표 2-16〉 3개 권역별 아동 연령별 분포

(단위: %)

3개 권역	계	0세	1세	2세	3세	4세	5세	6세
계	100.00	11.72	12.04	13.49	14.34	15.98	15.51	16.92
대도시	36.56	36.56	40.31	36.21	36.04	35.50	35.16	36.87
중소도시	37.16	38.17	33.51	39.25	37.14	35.90	39.02	36.87
읍 면	26.28	25.27	26.18	24.53	26.81	28.60	25.81	26.26

제3장 2009년 전국 보육실태조사 표본설계

제1절 2009년 표본설계의 특징

2004년 보육교육실태조사와는 달리 2009년의 조사에서는 보육실태만을 조사하기 때문에 전국의 0~6세 아동의 분포 따라 층별로 표본을 배분하도록 한다.

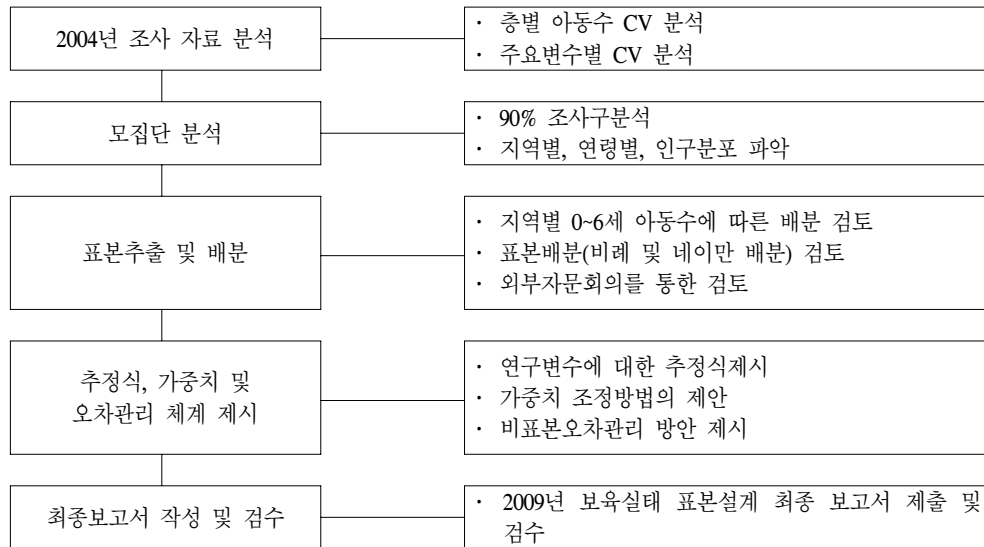
또한 지역별 층화 기준은 2004년과 같이 전국을 3개 층으로 구분하고, 조사구 유형에 따라 정렬한 후 약 250개 조사구를 표본으로 추출하도록 한다.

2009년의 조사에서는 인구 총 조사 90% 전수 조사구를 이용하며, 지역별 아동 및 보육시설에 관한 정보를 이용하여 층별 표본을 배분하도록 하며, 조사현장에서 발생할 수 있는 조사거부, 재개발, 또는 천재지변 등과 더불어 지역별 아동 출현율을 감안하여 약 10% 정도의 예비 표본을 제공한다.

또한 표본가구의 규모는 250개 조사구내 평균 60가구를 조사할 경우 약 15,000가구를 목표 가구수로 설정하고, 조사 완수율은 약 85%로 산정할 경우 12,750가구가 될 수 있도록 한다.

이와 같은 내용을 수행하기위해 ① 모집단 분석, ② 표본설계(배분 및 추출), ③ 자문회의를 통한 최종방안 확정, ④ 표본지역별 가구명부 및 요도 복사, ⑤ 표본지역 확정을 위한 조사원 교육, ⑥ 최종 표본지역 확정, ⑦ 표본설계 보고서 작성 및 제출 등의 절차에 따라 과업을 수행하고자 한다.

[그림 3-1] 2009년 보육실태조사 표본설계 절차



제2절 모집단 분석

1. 모집단 분석

새로운 표본설계의 모집단조사구는 2005년도에 실시한 인구주택총조사에서 전수 조사가 실시된 90% 전수조사구 238,585개로 구성되었다. 표본조사가 실시된 10% 조사구는 통계청에서 실시되고 있는 각종 조사에 사용되고 있어 사용이 제한된다.

새로운 표본설계에서는 섬조사구(903개)를 제외한 237,682개의 아파트조사구와 일반조사구만을 모집단조사구로 사용한다. 이는 아파트 조사구와 일반 조사구를 제외한 조사구에서의 조사가 현실적으로 어렵기 때문이다. <표 3-1>에 의하면, 아파트조사구와 일반조사구는 전체 조사구의 99.62%를 차지하고 있으며, 가구수와 인구수도 각각 전체의 99.64%와 99.72%로 거의 대부분을 차지하고 있다.

〈표 3-1〉 조사구 특성별 분포

(단위: 개, 명, %)

구분	합계	아파트	일반	섬
조사구수 (비율)	238,585 (100.00)	95,728 (40.12)	141,954 (59.50)	903 (0.38)
가구수 (비율)	14,312,048 (100.00)	5,881,924 (41.10)	8,378,179 (58.54)	51,945 (0.36)
인구수 (비율)	41,312,775 (100.00)	19,029,478 (46.06)	22,167,307 (53.66)	115,990 (0.28)

한편, <표 3-2>로부터 거처종류별로 보면 단독주택이 전체의 44.10%인 6,311,694가구이고 아파트가 6,005,176가구로 전체의 41.96%를 차지하고 있어 단독주택과 아파트가 전체가구의 86.06%가 된다.

〈표 3-2〉 거처 종류별 가구수

(단위: 개, %)

구분	합계	단독주택	아파트	연립주택	다세대 주택	영업용 건물 내 주택	주택이외 거처
가구수	14,312,048	6,311,694	6,005,176	477,679	1,063,522	254,436	199,541
비율	100.00	44.10	41.96	3.34	7.43	1.78	1.39

<표 3-3>에 의하면, 전체 조사구에서 동부가 191,330개로 전체의 80.19%이고 읍면부가 47,256개로 19.81%이며, 가구수도 조사구수와 거의 같은 비율로 분포하고 있다. 그러나 동부의 인구수는 조사구나 가구수에 비해 좀 더 많은 82.24%를 차지하고 있다. 전체 조사구 중에서 서울의 조사구는 20.72%이고 경기도가 20.87%로 두 지역이 전체 조사구의 41.59%나 된다. 가구수와 인구수도 비슷한 정도를 차지하고 있다. 상세한 지역별 동부, 읍면부 조사구와 가구 및 인구 분포는 <표 3-3>에 있다.

〈표 3-3〉 지역별 조사구, 가구 및 인구 분포

(단위: 개, 명, %)

지역	전체			동부			읍면부		
	조사구수	가구수	인구수	조사구수	가구수	인구수	조사구수	가구수	인구수
총계 (비율)	238,585 (100.00)	14,312,048 (100.00)	41,312,775 (100.00)	191,330 (80.19)	11,570,167 (80.84)	33,974,399 (82.24)	47,256 (19.81)	2,741,881 (19.16)	7,338,376 (17.76)
서울	49,440	3,005,559	8,667,100	49,440	3,005,559	8,667,100	-	-	-
부산	18,179	1,075,332	3,107,992	17,773	1,053,154	3,044,049	406	2,2178	63,943
대구	12,128	738,953	2,179,663	11,413	696,314	2,048,122	715	42,639	131,541
인천	12,814	743,581	2,235,126	12,452	722,165	2,179,179	362	21,416	55,947
광주	6,793	416,208	1,241,582	6,793	416,208	1,241,582	-	-	-
대전	7,268	434,710	1,269,860	7,268	434,710	1,269,860	-	-	-
울산	5,053	308,537	925,705	4,248	260,251	781,398	805	48,286	144,307
경기	49,804	3,021,968	9,167,635	40,786	2,505,031	7,619,860	9,018	516,937	1,547,775
강원	7,871	462,601	1,253,121	4,790	285,326	783,756	3,081	177,275	469,365
충북	7,499	450,291	1,253,728	4,550	279,771	797,778	2,949	170,520	455,950
충남	9,891	589,435	1,607,199	3,498	216,568	621,938	6,393	372,867	985,261
전북	9,297	545,531	1,521,691	6,204	369,987	1,087,347	3,093	175,544	434,344
전남	9,887	582,815	1,548,644	4,019	235,716	696,112	5,867	347,098	852,532
경북	14,025	830,597	2,209,958	6,980	419,252	1,175,917	7,045	411,345	1,034,041
경남	15,932	944,113	2,655,637	9,188	553,502	1,617,201	6,744	390,611	1,038,436
제주	2,704	161,818	468,134	1,928	116,653	343,200	776	45,165	124,934

<표 3-4>에서 일반 및 아파트 조사구만을 고려할 때 동부의 조사구는 191,286개로 전체 조사구의 80.48%를 차지하여 읍면부(19.52%)의 4.1배가 된다. 동부는 도시지역으로 전체 아파트조사구의 88.16%나 된다. 반면에 읍면부는 상대적으로 일반조사구의 비중이 높다. 서울, 광주, 그리고 대전에는 읍면부가 없고 모두 동부로 구성되어 있으며, 나머지 광역시에는 약간의 읍면부가 있다. 대부분의 지역이 동부의 조사구가 읍면부의 조사구보다 많다. 그러나 충남, 전남과 경북은 읍면부의 조사구가 동부의 조사구보다 많다. 한편 일반 및 아파트 조사구에 대한 가구수와 인구수에 대한 통계는 <표 3-4>, <표 3-5> 그리고 <표 3-6>에 제시하였다.

〈표 3-4〉 지역별 조사구 분포(아파트, 일반)

(단위: 개, %)

지역	전체			동부			읍면부		
	계	일반	아파트	계	일반	아파트	계	일반	아파트
총계 (비율)	237,682 (100.00)	141,954 (100.00)	95,728 (100.00)	191,286 (80.48)	107,066 (75.42)	84,220 (88.16)	46,396 (19.52)	34,888 (24.58)	11,508 (11.84)
서울	49,440	32,243	17,197	49,440	32,243	17,197	-	-	-
부산	18,165	10,817	7,348	17,759	10,557	7,202	406	260	146
대구	12,128	7,070	5,058	11,413	6,723	4,690	715	347	368
인천	12,713	7,076	5,637	12,449	6,839	5,610	264	237	27
광주	6,793	2,982	3,811	6,793	2,982	3,811	-	-	-
대전	7,268	3,812	3,456	7,268	3,812	3,456	-	-	-
울산	5,053	2,698	2,355	4,248	2,281	1,967	805	417	388
경기	49,803	26,311	23,492	40,786	20,519	20,267	9,017	5,792	3,225
강원	7,871	4,852	3,019	4,790	2,405	2,385	3,081	2,447	634
충북	7,499	4,457	3,042	4,550	2,182	2,368	2,949	2,275	674
충남	9,869	6,504	3,365	3,498	1,629	1,869	6,371	4,875	1,496
전북	9,275	5,545	3,730	6,204	2,751	3,453	3,071	2,794	277
전남	9,243	6,301	2,942	4,001	1,818	2,183	5,242	4,483	759
경북	14,025	9,529	4,496	6,980	3,846	3,134	7,045	5,683	1,362
경남	15,857	9,561	6,296	9,179	5,028	4,151	6,678	4,533	2,145
제주	2,680	2,196	484	1,928	1,451	477	752	745	7

* 총 조사구 237,682개는 아파트 조사구와 일반 조사구만을 합한 것이다.

〈표 3-5〉 지역별 가구 분포(아파트, 일반)

(단위: 개, %)

지역	전체			동부			읍면부		
	계	일반	아파트	계	일반	아파트	계	일반	아파트
총계 (비율)	14,260,103 (100.00)	8,378,179 (100.00)	5,881,924 (100.00)	11,567,849 (81.12)	6,360,644 (75.92)	5,207,205 (88.53)	2,692,254 (18.88)	2,017,535 (24.08)	674,719 (11.12)
서울	3,005,559	1,937,041	1,068,518	3,005,559	1,937,041	1,068,518	-	-	-
부산	1,074,589	623,668	450,921	1,052,411	610,218	442,193	22,178	13,450	8,728
대구	738,953	422,825	316,128	696,314	403,510	292,804	42,639	19,315	23,324
인천	737,976	399,023	338,953	722,054	384,745	337,309	15,922	14,278	1,644
광주	416,208	180,490	235,718	416,208	180,490	235,718	-	-	-
대전	434,710	224,936	209,774	434,710	224,936	209,774	-	-	-
울산	308,537	162,328	146,209	260,251	138,114	122,137	48,286	24,214	24,072
경기	3,021,932	1,547,053	1,474,879	2,505,031	1,225,549	1,279,482	516,901	321,504	195,397
강원	462,601	286,812	175,789	285,326	143,046	142,280	177,275	143,766	33,509
충북	450,291	267,315	182,976	279,771	134,390	145,381	170,520	132,925	37,595
충남	588,216	385,980	202,236	216,568	100,425	116,143	371,648	285,555	86,093
전북	544,563	317,258	227,305	369,987	158,310	211,677	174,576	158,948	15,628
전남	544,741	367,739	177,002	234,729	101,893	132,836	310,012	265,846	44,166
경북	830,597	558,120	272,477	419,252	226,562	192,690	411,345	331,558	79,787
경남	940,076	565,284	374,792	553,025	302,719	250,306	387,051	262,565	124,486
제주	160,554	132,307	28,247	116,653	88,696	27,957	43,901	43,611	290

〈표 3-6〉 지역별 인구 분포(아파트, 일반)

(단위: 명, %)

지역	전체			동부			읍면부		
	계	일반	아파트	계	일반	아파트	계	일반	아파트
총계 (비율)	41,196,785 (100.00)	22,167,307 (100.00)	19,029,478 (100.00)	33,969,281 (82.46)	17,053,034 (76.93)	16,916,247 (88.89)	7,227,504 (17.54)	5,114,273 (23.07)	2,113,231 (11.11)
서울	8,667,100	5,217,258	3,449,842	8,667,100	5,217,258	3,449,842	-	-	-
부산	3,106,319	1,689,256	1,417,063	3,042,376	1,652,134	1,390,242	63,943	37,122	26,821
대구	2,179,663	1,154,791	1,024,872	2,048,122	1,100,392	947,730	131,541	54,399	77,142
인천	2,222,205	1,097,671	1,124,534	2,178,971	1,059,677	1,119,294	43,234	37,994	5,240
광주	1,241,582	473,029	768,553	1,241,582	473,029	768,553	-	-	-
대전	1,269,860	589,387	680,473	1,269,860	589,387	680,473	-	-	-
울산	925,705	438,407	487,298	781,398	372,773	408,625	144,307	65,634	78,673
경기	9,167,570	4,237,392	4,930,178	7,619,860	3,321,320	4,298,540	1,547,710	916,072	631,638
강원	1,253,121	726,313	526,808	783,756	353,343	430,413	469,365	372,970	96,395
충북	1,253,728	687,573	566,155	797,778	344,051	453,727	455,950	343,522	112,428
충남	1,604,522	981,067	623,455	621,938	251,696	370,242	982,584	729,371	253,213
전북	1,519,141	801,930	717,211	1,087,347	417,536	669,811	431,794	384,394	47,400
전남	1,463,707	907,569	556,138	693,958	275,501	418,457	769,749	632,068	137,681
경북	2,209,958	1,355,008	854,950	1,175,917	568,896	607,021	1,034,041	786,112	247,929
경남	2,647,449	1,435,516	1,211,933	1,616,118	801,943	814,175	1,031,331	633,573	397,758
제주	465,155	375,140	90,015	343,200	254,098	89,102	121,955	121,042	913

〈표 3-7〉 지역별/조사구유형별 만 0~6세 아동 인구분포(90%자료)

(단위: 명)

지 역	계	동부			읍면부		
		계	일반	아파트	계	일반	아파트
전 국	3,241,075	2,687,719	1,070,397	1,617,322	553,356	256,534	296,822
서 울	589,468	589,468	318,036	271,432	0	-	-
부 산	194,396	189,330	83,349	105,981	5,066	2,051	3,015
대 구	160,496	147,601	64,513	83,088	12,895	3,088	9,807
인 천	176,781	174,231	72,889	101,342	2,550	1,875	675
광 주	108,583	108,583	27,861	80,722	0	-	-
대 전	104,801	104,801	37,891	66,910	0	-	-
울 산	79,828	64,728	23,852	40,876	15,100	3,631	11,469
경 기	832,953	689,279	244,617	444,662	143,674	56,771	86,903
강 원	96,849	62,935	18,383	44,552	33,914	20,150	13,764
충 북	99,530	67,239	19,399	47,840	32,291	16,720	15,571
충 남	129,724	59,265	14,043	45,222	70,459	33,436	37,023
전 북	119,361	94,902	22,995	71,907	24,459	17,039	7,420
전 남	116,337	64,066	15,331	48,735	52,271	30,659	21,612
경 북	167,020	101,736	33,482	68,254	65,284	31,747	33,537
경 남	221,538	135,628	50,359	85,269	85,910	30,008	55,902
제 주	43,410	33,927	23,397	10,530	9,483	9,359	124

본 연구에서 모집단으로 활용할 2005년 인구주택 총조사 90% 자료의 만 0~6세 아동 인구수를 분석한 결과는 <표 3-7>과 같다. 전체적으로 서울지역이 18.2%, 경기지역이 25.7%로 나타나고 있으며, 동부의 경우 일반조사구에서는 서울이 29.7% 경기지역이 22.9%로 분석되었다. 또한 읍면부의 경우 일반조사구에서는 경기지역이 22.1%, 충남지역이 13%의 순으로 나타나고 있으며, 아파트조사구에서는 경기지역이 29.3%, 경남지역이 18.8%의 순으로 분석되었다.

이러한 분포로부터 2005년 당시 0~6세 아동 수에 비례하여 표본을 배분할 경우 동부의 경우 서울과 경기지역에 많은 수의 표본이 배분될 것이며, 읍면부의 경우 경기지역과 경남 및 충남 지역에 많이 배분될 것으로 예상된다.

〈표 3-8〉 지역별/조사구유형별 0~6세 아동 연령의 CV분포(90%자료)

(단위: %)

지 역	동부		읍면부	
	일반	아파트	일반	아파트
서 울	63.31	62.26	-	-
부 산	59.92	60.21	58.49	62.60
대 구	59.93	59.68	59.15	61.44
인 천	61.18	59.72	57.90	62.04
광 주	58.67	60.55	-	-
대 전	62.26	60.30	-	-
울 산	60.92	60.19	56.70	60.50
경 기	62.23	59.98	58.45	62.35
강 원	59.64	60.31	57.81	64.62
충 북	60.15	60.69	56.88	63.46
충 남	60.00	61.14	57.81	62.83
전 북	57.68	60.46	56.80	63.06
전 남	57.44	59.36	55.78	61.83
경 북	59.63	60.68	57.50	61.59
경 남	60.51	60.01	57.12	62.17
제 주	61.58	59.42	58.20	63.97

<표 3-8>은 지역별/조사구 유형별 아동 연령의 CV 분포를 나타내고 있다. 0~6세 아동의 지역별 CV는 거의 대동소이한 값을 가지는 것으로 분석되고 있으나, 서울 동부의 일반 조사구의 경우 63.61%로 가장 크게 나타나고 있으며, 전남 동부의 일반조사구의 경우 57.44%로 가장 낮게 나타나고 있다. 읍면부의 경우 일반조사구에서는 전남지역이 55.78%로 가장 낮으며, 강원의 아파트조사구의 경우 64.62%로 가장 높게 나타났다. 이와 같은 분석으로부터 단순히 아동수에 비례하여 배분할 경우 <표 3-7>로부터 서울과 경기지역에 많은 표본이 배분될 것으로 예상되나, <표 3-8>의 아동 연령 CV의 분포로부터는 동부의 일반은 서울 및 대전, 아파트는 서울 및 경북, 읍면부의 일반조사구는 대구, 아파트는 제주 등일 것으로 예상된다.

한편 전국을 5개 권역으로 층화하여 각 권역별 아동 수 및 연령 CV를 분석한 결과가 <표 3-9>와 같다. 권역별로 아동수의 분포를 살펴보면 기타시도의 동부가 전체의 40.4%로 가장 높고, 다음으로 광역시의 동부가 24.4%를 차지하는 것으로 나타났다. 이와는 대조적으로 읍부와 면부의 경우 각각 10.1%와 7.0%로 매우 낮은 분포를 보이고 있어 권역별로 아동수에 비례하여 표본을 배분할 경우 읍부와 면부가 상대적으로 적은 표본수가 배분될 것으로 예상된다. 한편 권역별 CV분포를 살펴보면, 권역별 CV의 변동이 그다지 크지 크기 때문에 전체적으로 고른 배분이 이루어 질 것으로 예상되며, 아파트 조사의 경우 상대적으로 읍부에 비해 면부의 CV값이 커서 면부에 많은 수의 표본이 배분 될 것으로 예상된다.

<표 3-9> 5개 권역별/조사구유형별 0~6세 아동 수 및 CV분포(90%자료)

(단위: 명, %)

권역별	아동 수			CV	
	계	일반	아파트	일반	아파트
서울의 동	589,468	318,036	271,432	63.31	62.26
광역시의 동	789,274	310,355	478,919	60.47	60.08
기타시도의 동	1,308,977	442,006	866,971	61.13	60.15
읍부	326,347	115,146	211,201	57.67	61.59
면부	227,009	141,388	85,621	57.33	64.17

다음으로 전국을 3개 권역으로 층화하여 각 층별 0~6세 아동수와 연령 CV를 분석한 결과가 <표 3-10>과 같다. 대도시의 경우 전체의 42.5%를 차지하고, 다음으로 중소도시가 40.4%, 읍면지역이 17.1%를 차지하는 것으로 나타나, 층별 비례배분이 이루어 질 경우 대도시, 중소도시, 읍면의 순으로 표본이 배분될 것으로 예상된다. 한편, 연령 CV를 기준으로 살펴보면, 읍면지역의 아파트 지역의 CV가 가장 커서 해당 층에 상대적으로 많은 표본이 배분되며, 반대로 읍면지역의 일반조사구가 가장 적은 표본수가 배분될 것으로 예상된다.

〈표 3-10〉 3개 권역별/조사구유형별 0~6세 아동 수 및 CV분포(90%자료)
(단위: 명, %)

권역별	아동 수		CV	
	일반	아파트	일반	아파트
대도시	628,391	750,351	61.91	60.87
중소도시	442,006	866,971	61.13	60.15
읍면	256,534	296,822	57.48	62.34

앞의 분석과는 별도로 2005년 인구주택총조사 결과와 2009년 현재와의 인구변동성을 감안하여 2008년과 2009년 인구추계값을 이용하여 연령별 분포를 파악하는 것이 타당할 것이다. 그러나 2009년 현재 인구추계의 경우 자료의 제한으로 지역별, 연령대별 분포를 파악할 수 없다. 따라서 2008년과 2009년 만 0~4세의 지역별, 성별 인구분포만을 이용하여 지역별, 성별 인구분포를 파악한 결과 <표 3-11>과 같다.

한편 <표 3-11>로부터 만 0~4세까지 연령별, 성별 추계인구를 파악해보면, 2008년의 경우 총 추계인구수 48,606,787명 중 2,255,184명으로 전체의 4.63%로 나타나며, 2009년의 경우 총 추계인구 48,746,693명중 2,216,919명으로 전체의 4.55%로 추계되고 있다. 이러한 근거를 바탕으로 전국의 만 0~6세 이하의 아동 출현율은 약 5.5% 정도가 될 것이며, 전국 출산율이 2005년 당시에 비해 약간 상승한 결과를 반영하더라도 5.5~6.5% 정도가 될 것으로 예상된다. 따라서 위와 같은 아동 출현율을 바탕으로 4인가구를 기준으로 전국 15,000가구를 조사할 경우 0~6세 이하의 조사대상 아동은 약 3,300~3,900명 정도로 예상된다.

결과적으로 지역별로 조사구 규모에 따라 표본 조사구를 배분하기 보다는 지역별 아동인구수에 따른 표본을 배분하는 방안을 고려하는 것이 바람직 할 것이다.

〈표 3-11〉 지역별, 성별 만 0~4세 추계 인구분포

(단위: 명)

지 역	2008년			2009년		
	계	남	여	계	남	여
전 국	2,255,184	1,169,845	1,085,339	2,216,919	1,149,300	1,067,619
서 울	438,910	226,436	212,474	442,136	227,867	214,269
부 산	128,418	66,602	61,816	125,337	65,044	60,293
대 구	103,850	54,687	49,163	100,791	53,039	47,752
인 천	122,971	63,374	59,597	121,289	62,576	58,713
광 주	72,756	37,595	35,161	71,639	37,064	34,575
대 전	74,516	38,654	35,862	73,598	38,146	35,452
울 산	53,966	28,526	25,440	52,686	27,796	24,890
경 기	601,860	310,579	291,281	599,336	309,394	289,942
강 원	62,969	32,738	30,231	59,719	31,041	28,678
충 북	67,625	35,208	32,417	65,436	33,970	31,466
충 남	91,703	47,476	44,227	89,397	46,215	43,182
전 북	74,767	38,653	36,114	70,042	36,207	33,835
전 남	73,509	38,164	35,345	68,700	35,565	33,135
경 북	110,039	57,742	52,297	106,012	55,498	50,514
경 남	148,159	78,037	70,122	142,700	75,095	67,605
제 주	29,166	15,374	13,792	28,101	14,783	13,318

자료: KOSIS, 2008년, 2009년 연령별 추계인구수, 통계청.

2. 지역별 보육시설 분포

지역별 아동수 분포와 더불어 추가적으로 지역별 영유아 보육시설 분포를 참조하여 지역별 표본을 배분하는 방안을 검토하도록 한다. 이와 같이 지역별 보육시설 규모를 고려한 이유는 영유아가 많은 지역이 보육서비스의 수요가 많고, 따라서 보육시설 규모를 함께 고려할 경우 단순한 인구비례에 의한 표본배분으로 인한 문제점을 보완할 수 있기 때문이다. 그러나 본 연구에서는 특광역시별로 지역을 층화하여 조사하지 않기 때문에 지역별 보육시설의 분포는 단순히 참고자료로만 이용하였다.

〈표 3-12〉 지역별, 설립주체별 보육시설 현황(2008년 6월말 현재)

(단위 : 개소)

구분	계	설립주체별							
		국·공립 보육시설	법인 보육시설	민간개인시설			부모협동 보육시설	가정 보육시설	직장 보육시설
				소계	법인의외	민간개인			
총 계	32,149	1,769	1,482	14,127	973	13,154	61	14,375	335
서울	5,519	622	35	2,562	224	2,338	16	2,198	86
부산	1,618	131	87	857	39	818	3	521	19
대구	1,339	28	124	712	39	673	4	458	13
인천	1,608	71	9	746	38	708	2	757	23
광주	1,057	22	113	366	28	338	1	543	12
대전	1,263	28	42	428	26	402	1	748	16
울산	587	28	13	352	6	346	0	184	10
경기	9,233	366	83	3,590	147	3,443	30	5,084	80
강원	852	73	116	409	51	358	2	246	6
충북	897	41	106	426	36	390	1	313	10
충남	1,289	50	135	586	66	520	1	502	15
전북	1,426	40	164	544	82	462	0	674	4
전남	1,024	54	181	411	56	355	0	369	9
경북	1,623	100	88	833	40	793	0	587	15
경남	2,369	99	109	1,040	55	985	0	1,108	13
제주	445	16	77	265	40	225	0	83	4

자료: 보건복지가족부

또한 지역별 연령별 인구분포와 함께, 지역별 보육시설 분포를 동시에 고려함으로써 인구수 비례에 의한 표본 설계시 모집단으로 활용하고 있는 2005년 인구주택총조사 자료는 2009년 시점의 영유아 아동의 분포를 참고할 수 없기 때문에 연구대상 집단인 0~6세 인구의 분포를 적절히 설명할 수 있는 보조 자료로서 지역별 보육시설 분포를 참고로 검토하도록 하였다.

〈표 3-13〉 지역별, 도시규모별 보육시설 현황(2008년 6월말 현재)

(단위: 개소)

구 분	계	설립주체별						
		국·공립 보육시설	법인 보육시설	민간보육시설		부모 협동	가정 보육시설	직장 보육시설
				법인의외	민간개인			
전국 총계	32,149	1,769	1,482	973	13,154	61	14,375	335
대도시	12,681	908	391	394	5,464	27	5,324	173
중소도시	13,486	454	342	219	5,165	32	7,151	123
농어촌	5,982	407	749	360	2,525	2	1,900	39

이와 더불어 <표 3-12>의 지역별 보육시설 현황을 <표 3-13>과 같이 도시규모별로 살펴보면 대도시지역의 보육시설은 12,681개소로 전체의 39.44%, 중소도시는 13,468개소로 41.89% 농어촌은 5,986개소로 18.61%를 차지하는 것으로 파악되었다. 각 지역별 보육시설 유형에 따른 표본배분방법은 지역별, 도시규모별로 보육시설수가 0인 지역이 있으므로 16개 광역단체별로 표본배분은 조사의 목적과 부합되지 않기 때문에 보다 큰 규모로서 2004년에 고려한 지역별 3개 권역인 특·광역시의 동부, 중소도시의 동부, 기타지역의 읍면부를 고려하거나, 대도시, 중소도시, 농어촌과 같은 도시규모에 따른 표본배분을 고려할 수 있을 것이다.

제3절 층화(Stratification)

1. 층화 기준

앞에서 언급한 바와 같이 단순히 지역별 또는 3개 층별 조사구 규모에 비례하여 표본을 배분할 경우 아동 출현율이 현격히 떨어질 가능성이 있으며, 이를 보완하기 위해 지역별 또는 3개 층 또는 도시규모별 표본배분에 지역별 0~6세 이하의 영유아 분포를 감안한 표본배분을 고려할 수 있을 것이다.

또한 이와 더불어 도시규모별 보육 시설 분포를 보조 자료로 활용하여 최종 표본 배분에 이용함으로써 도시규모별 아동 출현율을 보다 엄밀히 반영할 수 있을 것으로 판단된다.

따라서 표본 조사구에 대한 구체적인 표본배분 방안으로는 첫째, 도시규모별 조사구 규모에 따른 비례배분을 고려할 수 있다. 둘째, 도시규모별 0~6세까지의 영유아 아동의 분포를 고려한 지역별 표본배분 방안을 고려할 수 있다. 셋째, 도시규모별 0~6세까지의 영유아 아동 분포와 지역별 보육시설 규모를 감안한 표본배분 방법을 고려할 수 있다.

2004년의 경우 서울지역에서 30개, 광역시의 동지역에서 51개, 기타시도 동지역에서 71개 조사구를 배분하고, 읍면지역에 각각 32개와 6개의 조사구를 배분하였다. 이때, 고려된 사항은 지역별 조사구 규모 및 지역별 추출률로서 별도의 아동 인구 및 보육시설에 관한 사항은 고려되지 않았다.

2009년 표본설계에서는 2004년과 같이 층화방법으로 (1) 전국을 5개 권역으로 층화하는 방법과 (2) 전국을 3개 권역으로 층화하는 방법을 검토하고자 한다. 한편 표본배분은 지역별 아동 인구수와 보육시설 규모에 따른 표본 배분을 고려한다. 이때, 비례배분 또는 네이만 배분 등의 표본배분 방식을 함께 고려하도록 한다.

2. 층화

효율적인 모집단의 층화를 위해서는 층화변수를 선택하는 동시에 각 층의 분산을 최소화할 수 있도록 층의 경계점을 복합적으로 결정해야 한다. 층화 변수가 성별 또는 지역과 같은 범주형 변수인 경우 별도의 층화 경계점을 별도로 결정할 필요가 없으나, 만일 층화변수값이 연속형인 경우 반드시 최적의 층화경계점을 결정하도록 해야 한다. 본 연구에서는 모집단으로 활용될 자료가 인구센서스 자료이고, 이때 층화변수로 고려할 조사변수가 제한적이기 때문에 별도의 최적 층화 경계점을 결정하는 것은 한계가 있다. 따라서 2장에서 다룬바 있는 지역별 층화방법을 검토하여 본 연구에서 주요 추정사항인 아동 보육 수요율 추정값이 안정적으로 산출될 수 있도록

록 층화를 결정하고자 한다.

먼저 2장에서 검토한 몇 가지 주요 분석 결과로부터 최종방안으로 고려된 층화방법은 (1) 전국을 5개 권역으로 층화하는 방법과 (2) 전국을 3개 권역으로 층화하는 방법을 검토하고자 한다.

가. 5개 권역으로의 층화

전국을 동부와 읍면부로 구분하고, 이를 다시 “서울의 동부”, “대도시의 동부”, “기타시의 동부”, “읍부”, “면부” 의 5개 권역으로 층화하는 방법이다. 표본수 및 영유아 아동 출현율을 감안할 때 16개 특광역시로 층화하는 방법에 비해 안정적인 통계산출이 가능하다는 장점이 있으나, “읍부” 와 “면부” 를 별도로 구분할 경우 해당 지역의 아동출현율이 상대적으로 낮아 안정적인 통계산출이 곤란할 수 있기 때문에 특별한 주의가 필요할 것으로 판단된다.

물론 2004년 조사결과와 공표 방법과 같이 5개 권역으로 층화한 후 결과산출시 다시 3개 권역으로 재분류하여 발표하는 방법도 고려할 수 있는 방안이라고 판단된다.

나. 3개 권역으로의 층화

한편 2004년과는 달리 표본 설계시 전국을 “대도시”, “중소도시”, “농어촌(읍면부)” 의 3개 권역으로 층화하여 표본을 추출하는 방법을 고려할 수 있다. 이와 같이 큰 권역으로 층화할 경우 층별 표본수가 안정적으로 확보될 수 있다는 장점과 더불어 영유아 아동출현율의 안정을 확보할 수 있다는 장점이 있는 반면, 농어촌(읍면부) 층이 도농 복합지역에서는 도시지역의 특성을 갖기 때문에 농어촌지역의 특성을 적절히 반영하기 어려운 문제점이 있다. 그러나 전국을 고려할 때 도농 복합 지역의 수가 그다지 많지 않기 때문에 전국 추정에서는 영향성이 미미할 것으로 예상된다.

결과적으로 2장의 2004년 조사결과와 앞에서 제시한 2가지 권역별 층화방법의 장단점을 면밀히 검토한 결과 모집단을 3개 권역별로 층화하여 각 층별로 최종 표본수를 결정하고, 최종 통계결과 역시 3개 권역별로 산출하도록 하였다.

제4절 표본 규모의 결정 및 배분

1. 표본규모의 결정

표본조사에서 모집단의 분포를 적절히 반영할 수 있는 적정 수준의 표본크기를 결정하는 것은 매우 중요하다. 과대하게 큰 규모의 표본수를 고려할 경우 조사비용 및 비표본오차의 발생으로 조사의 효율성이 떨어지게 되며, 반대로 너무 적은 규모의 표본수를 결정할 경우 표본 오차가 매우 커짐으로서 통계값의 대표성에 문제가 되기 때문이다. 따라서 제한된 조사비용과 여건을 적절히 반영하고, 또한 표본오차를 적절히 통제할 수 있도록 표본수를 결정하는 것이 바람직 할 것이다. 한편 인구주택총조사 모집단 자료에는 집락에 대한 정보가 있기 때문에 실제로 표본추출단계에서는 표본가구를 직접 추출하기보다는 이들의 집합인 집락(조사구)을 1차 추출단위로 하고 있다. 따라서 최적의 표본수 결정에서는 먼저 1단계로 추출되는 집락의 최적 규모를 검토해야 한다. 즉, 1단계로 집락(조사구)의 규모를 결정하고, 2단계에서 표본가구의 규모를 결정하게 되는데, 이때 표본가구의 규모를 어느 정도로 할 것인가에 따라 추정효율의 정도가 변동하게 된다. 또한 조사의 수행과 표본의 관리 측면에서 표본가구의 규모를 어느 정도로 할 것인가에 따라 다르게 된다.

본 연구에서는 아동 출현율에 따른 보육수요율이 주요 분석 목적이므로 아동 출현율과 허용오차에 대한 조건하에서 표본수를 결정하도록 한다.

〈표 3-14〉 95% 신뢰수준 하에서 허용오차와 보육 수요율에 따른 표본수

허용오차(d)	보육수요율(p)	n
0.1	0.35	88
	0.40	93
0.07	0.35	179
	0.40	189
0.06	0.35	243
	0.40	257
0.05	0.35	350
	0.40	369
0.03	0.35	969
	0.40	1,021
0.02	0.35	2,166
	0.40	2,283
0.01	0.35	8,431
	0.40	8,876

2004년도 조사로부터 단기 보육수요율이 영아의 경우 25%, 유아의 경우 43.7%로 추계되어 전체적으로는 35.6%의 수요율로 추계되었다. 이러한 배경 하에서 최소한 2004년의 조사와 일관성을 유지하기 위해 2009년 조사의 경우 전체적인 보육수요율을 35~40%로 유지하기 위해 허용오차와 조사비용 등을 감안할 때 허용오차 6% 내에서 243~257개 조사구가 필요한 것으로 판단된다.

따라서 조사예산 및 아동 출현율 등을 감안할 경우 허용오차 6%내외로 보육수요율 추계치를 40%로 가정하여 전국적으로 약 257개의 표본조사구를 추출하여 최종적인 표본 가구수는 조사구당 평균 60가구를 가정할 경우 약 15,420가구를 조사하여 조사의 일관성을 유지하기로 한다.

2. 표본 배분

표본의 배분은 앞에서 언급한 층별로 전체 표본을 어떠한 기준에 의해 배분할 것인가를 결정하는 단계이다. 따라서 각 대안별로 층화기준에 따라 전체 257개 조사구를 배분하여 가장 적절한 배분 방안을 결정하기로 한다. 이때 층별 배분 기준은 크

게 두 가지로 고려할 수 있는데, (1) 층별 조사구 규모에 비례하도록 배분하는 배례 배분(proportional allocation)과 (2) 층별 변동에 따라 변동이 큰 층에 표본을 많이 배 분하는 네이만 배분(Neyman allocation)방법이다.

- 비례 배분 : $n_h = n \frac{N_h}{N}$
- 네이만 배분 : $n_h = n \frac{N_h S_h}{\sum_{h=1}^H N_h S_h}$ 또는 $n_h = n \frac{N_h \sqrt{P_h Q_h}}{\sum_{h=1}^H N_h \sqrt{P_h Q_h}}$

네이만 배분의 경우 기존의 조사로부터 층별 보육 수요율 추정값을 활용할 경우 두 번째 공식을 활용할 수 있다.

가. 5개 권역별 조사구 규모와 변동에 따른 표본배분

층별 모집단 조사구 수에 대한 비례배분을 적용하기위해 먼저 모집단의 권역별 조사구수에 따른 층별 표본조사구 배분 결과는 다음과 같다. <표 3-15>의 비례배분 결과 기타시의 동부에 89개로 가장 많은 조사구가 배분되었고, 다음은 광역시의 동 부가 65개, 서울의 동부가 53개 조사구가 배분되었다. 한편 읍부와 면부의 경우 각 각 22개와 28개 조사구가 배분되어 5개 권역별 배분이 조사구의 크기에 따라 배분 되었다. 한편 조사구 유형에 따라 일반 조사구와 아파트 조사구에 따른 표본배분 결 과를 살펴보면, 다른 지역에 비해 읍부와 면부의 아파트 조사구 배분결과 각각 9개 와 4개로서 조사구별 편차가 발생할 가능성이 있는 것으로 나타났다.

〈표 3-15〉 5개 권역별 조사구 규모에 따른 층별 비례배분 현황

(단위: 개)

권역	모집단 조사구			표본조사구(비례배분)		
	계	일반	아파트	계	일반	아파트
전국	237,682	141,954	95,728	257	153	104
서울의 동	49,440	32,243	17,197	53	35	18
광역시의 동	59,930	33,194	26,736	65	36	29
기타시의 동	81,916	41,629	40,287	89	45	44
읍	20,350	12,181	8,169	22	13	9
면	26,046	22,707	3,339	28	24	4

다음으로 층별 조사구 규모의 변동을 고려한 네이만 배분 결과가 <표 3-16>에 제시되었다. 전체적으로 비례배분보다는 지역별, 조사구별로 고르게 배분되었음을 알 수 있다. 즉, 비례배분의 경우 읍부와 면부의 아파트 조사구가 10개미만으로 배분된 반면, 네이만 배분의 경우 11개와 16개로 상대적으로 많은 조사구가 배분되어 지역별, 조사구 유형별 편차를 줄일 수 있을 것으로 판단된다.

〈표 3-16〉 5개 권역별 네이만 배분에 의한 층별 표본 배분 현황

(단위: 개)

권역	모집단 표준편차			표본조사구(네이만 배분)		
	계	일반	아파트	계	일반	아파트
전국	11.90	13.17	9.52	257	147	110
서울동	11.88	12.97	9.36	52	30	22
광역시동	10.74	11.95	8.78	59	34	25
기타시동	11.74	13.55	9.37	89	53	36
읍	13.35	14.93	10.51	25	14	11
면	13.20	13.28	12.62	32	16	16

나. 3개 권역별 조사구 규모와 변동에 따른 표본배분

3개 권역별 층화에 따른 층별 모집단 조사구수에 대한 비례배분을 적용하기 위해 먼저 모집단의 층별 조사구수에 따른 표본조사구 배분 결과는 다음과 같다.

〈표 3-17〉 3개 권역별 비례배분에 의한 층별 표본 배분 현황

(단위: 개)

권역	모집단 조사구			표본조사구(비례배분)		
	계	일반	아파트	계	일반	아파트
전국	237,682	141,954	95,728	257	153	104
대도시	109,370	65,437	43,933	118	71	48
중소도시	81,916	41,629	40,287	89	45	43
읍면	46,396	34,888	11,508	50	37	13

<표 3-17>로부터 비례배분결과 대도시 지역이 전체의 45.9%인 118개 조사구로 가장 많이 배분되었고, 다음으로 중소도시 층이 34.6%인 89개 조사구가 배분되었다. 읍면지역은 50개 조사구가 배분되었고, 특히 조사구 유형에 따른 배분 결과를 보면 일반조사구가 37개인 반면 아파트조사구는 13개 밖에 배분되지 않아 지역별로 큰 편차가 나고 있음을 알 수 있다. 이와 같은 맥락에서 대도시지역의 경우 일반조사구는 71개, 아파트 조사구는 48개로 읍면지역과 마찬가지로 아파트조사구의 수가 상대적으로 적게 배분되었다.

〈표 3-18〉 3개 권역별 네이만 배분에 의한 층별 표본 배분 현황

(단위: 개)

권역	모집단 표준편차			표본조사구(네이만 배분)		
	계	일반	아파트	계	일반	아파트
전국	11.90	13.17	9.52	257	173	84
대도시	11.28	12.49	9.02	113	76	37
중소도시	11.74	13.55	9.37	87	52	35
읍면	13.28	13.88	11.24	57	45	12

다음으로 모집단을 3개층으로 층화하여 각 층별 변동을 고려한 네이만 배분 결과가 <표 3-18>에 제시되었다. 3개 층에 대한 네이만 배분결과와 <표 3-17>에 제시된 비례배분결과가 거의 유사함을 알 수 있다. 단지 읍면지역의 일반 조사구의 경우 아파트 조사구에 비해 가구수 변동이 크게 나타남에 따라 비례배분 보다 해당 셀에 상대적으로 많은 수의 조사구가 배분되었다.

다. 5개 권역별 아동 규모와 연령 변동에 따른 표본배분

다음의 <표 3-19>는 5개 권역별 아동수에 비례하도록 257개 조사구를 배례배분한 결과이다. 일반조사구는 105개, 아파트조사구는 152개가 배분되었으며, 각 권역별로는 서울의 동부 47개 조사구가 일반조사구는 25개, 아파트조사구는 22개가 배분되었다. 광역시의 동부에서는 일반조사구는 25개 아파트조사구는 38개가 배분되었고, 기타시의 동부에서는 103개 조사구가 각각 35개와 68개 조사구로 배분되었다. 읍지역의 경우 전체 26개 조사구가 배분되어, 이중 9개는 일반조사구, 17개는 아파트조사구로 배분되었다. 면부는 18개 조사구가 일반 11개, 아파트조사구는 7개가 각각 배분되어 읍지역과는 달리 아파트조사구가 상대적으로 적게 배분되었다.

<표 3-19> 5개 권역별 아동수에 따른 층별 비례배분 현황

(단위: 개)

권역	모집단 아동 수(명)			표본조사구(비례배분)		
	계	일반	아파트	계	일반	아파트
전국	3,241,075	1,326,931	1,914,144	257	105	152
서울의 동	589,468	318,036	271,432	47	25	22
광역시의 동	789,274	310,355	478,919	63	25	38
기타시동	1,308,977	442,006	866,971	103	35	68
읍	326,347	115,146	211,201	26	9	17
면	227,009	141,388	85,621	18	11	7

한편 <표 3-20>은 권역별 아동 연령의 변동에 따라 257개 조사구를 네이만 배분한 결과이다. 권역별로 아동 연령의 변동이 적기 때문에 결과적으로 비례배분한 결과와 매우 유사함을 알 수 있다.

<표 3-20> 5개 권역별 아동 연령 변동에 따른 층별 네이만 배분 현황

(단위: 개)

권역	모집단 표준편차			표본조사구(네이만배분)		
	계	일반	아파트	계	일반	아파트
전국	1.98	1.98	1.98	257	105	152
서울의 동	1.97	2.00	1.99	46	25	21
광역시외의 동	1.97	1.99	1.98	63	25	38
기타시동	1.97	1.99	1.98	104	35	69
읍	1.97	1.97	1.97	26	9	17
면	1.97	1.96	1.98	18	11	7

라. 3개 권역별 아동 규모와 연령 변동에 따른 표본배분

다음의 <표 3-21>과 <표 3-22>는 전국을 3개 층으로 층화하여 층별 아동수 및 연령 변동에 따라 비례배분과 네이만 배분한 결과이다. 두 결과로 부터 비례배분결과와 네이만 배분결과가 거의 유사함을 알 수 있다. 이는 아동 연령의 변동이 거의 층별로 동일하기 때문이다.

<표 3-21> 3개 권역별 비례배분에 의한 층별 표본 배분 현황

(단위: 개)

권역	모집단 아동수(명)			표본조사구(비례배분)		
	계	일반	아파트	계	일반	아파트
전국	3,241,075	1,326,931	1,914,144	257	105	152
대도시	1,378,742	628,391	750,351	110	50	60
중소도시	1,308,977	442,006	866,971	103	35	68
읍면	553,356	256,534	296,822	44	20	24

〈표 3-22〉 3개 권역별 네이만 배분에 의한 층별 표본 배분 현황

(단위: 개)

권역	모집단 표준편차			표본조사구(네이만 배분)		
	계	일반	아파트	계	일반	아파트
전국	1.98	1.98	1.98	257	105	152
대도시	1.99	2.00	1.98	109	50	59
중소도시	1.99	1.99	1.98	104	35	69
읍면	1.97	1.96	1.98	44	20	24

최종적으로 앞에서 논의한 표본배분 방법은 <표 3-21>과 같이 0~6세 아동 출현을 등을 감안하여 층별로 아동 수에 따라 비례 배분하는 방법을 적용하며, 층화 방법은 2004년과 같이 3개 권역별 층화 방법을 적용하도록 하였다.

제5절 표본 추출방법

1. 표본조사구 추출

16개 광역시(광역시와 광역자치단체)를 “대도시”, “중소도시”, “읍면”의 3개 층으로 구분하여 각 층별로 표본 조사구를 추출할 때 층내를 “일반조사구”와 “아파트 조사구”로 나누어 계통추출법을 적용한다. 계통추출을 위해서 각 층내에서 조사구를 행정구역으로 정렬하고, 아파트조사구와 일반조사구 순으로 정돈한 후에 각각 아동수를 크기척도(MOS)로 고려하여 이에 확률비례하는 층화확률비례계통추출법을 사용한다. 이러한 추출방법은 표본 조사구를 층내에서 지역별로 고르게 뽑을 수 있고 또한 행정구역 내에서 아파트조사구와 일반조사구가 균형있게 추출되어서 표본의 대표성이 커진다. 모집단조사구는 2005년도 인구주택총조사 90% 자료를 사용하는 관계로 표본으로 추출된 일부의 조사구에 변동이 있을 수 있다. 표본으로 추출된 조사구를 조사에 사용할 수 없는 경우에는 <부록>에 주어진 예비 표본조사구를 사용한다.

2. 표본가구와 응답자 추출

표본조사를 위한 조사대상 가구는 표본조사구로부터 추출한다. 조사모집단 자료에 의하면, 조사구는 평균 60가구로 구성되어 있다. 『2009년 보육실태조사』에서는 표본 조사구에 있는 가구 중 60가구를 표본가구로 추출하도록 하였다. 만일 60가구 미만의 조사구에 대해서는 해당 조사구내 가구를 모두 조사 하도록 한다.

표본가구를 추출하기 위해서는 표본 조사구에 대한 가구목록(통계청으로부터 제공 받음)을 사용한다. 통계청에서 제공한 가구목록을 사용하기 전에 조사팀은 표본 조사구를 사전에 방문해서 가구목록과 실제 거주하고 있는 가구들을 확인한다. 이때 목록에 누락되었거나 추가되는 가구를 파악해서 가구목록을 새로이 작성하여 실제 조사에 사용한다.

표본가구추출은 비복원 단순확률추출법을 사용한다. 표본으로 추출된 조사구내의 가구목록을 작성한 후 난수를 발생하여, 발생된 난수에 해당되는 가구를 표본가구로 한다. 예를 들어 표본조사구내의 가구가 60가구 이상일 경우, EXCEL, SAS 등을 이용하여 1~60사이의 비복원 난수를 60개를 발생하여 번호에 해당되는 가구를 표본가구로 한다. 한편 표본가구 중에서 표본으로 사용할 수 없는 가구가 있을 경우에는 교체표본을 사용하는 데, 이를 위해서 예비표본가구를 선정한다. 예비표본가구는 주어진 표본 조사구에서 표본가구로 선정된 가구를 제외하고 나머지 가구들을 대상으로 필요한 수의 예비표본을 표본가구의 선정과 동일한 비복원 단순확률추출법을 사용한다.

제4장 추정

제1절 기본 가중치 산정

일반적으로 추정단계에서 가중치를 이용하면 모집단에 대한 특성치인 모수에 대한 비편향추정량(unbiased estimator)을 얻을 수 있다. 만약 통계분석 과정에서 가중치를 무시하고 분석한 추정치는 심각한 편향이 발생할 수 있다. 표본의 크기가 큰 대규모 조사에서 문제가 되는 것은 추정량의 편향이기 때문에 추정과정에서 반드시 가중치를 이용해야 한다. 일반적으로 복합표본조사의 가중치는 설계 가중치, 무응답에 대한 조정, 사후층화에 대한 조정 등의 세 가지 요인을 통합하여 산정된다.

본 조사에서는 전국의 행정구역을 “대도시”, “중소도시”, “읍면”지역의 3개 권역으로 층화하고 각 층내에서 조사구 유형에 따라 일반조사구와 아파트조사구로 구분하였다. 2009년 보육실태조사의 가중치는 각 층에서 모집단 가구 수와 표본의 가구 수의 복원배율로 계산된다.

먼저 본 조사의 표본추출은 2단계로 수행되기 때문에 1단계에서는 조사구, 2단계에서는 가구 에 대한 추출확률을 계산해야 한다. 따라서 각 층별로 배분된 조사구 규모에 따라 모집단 층으로부터 추출된 단위들의 추출확률(p_{hij} : sampling probability)은 다음과 같이 계산된다.

$$p_{hij} = p_{hi} \times p_{hj(i)} = \frac{n_h}{N_h} \times \frac{m_{hi}}{M_{hi}} = \frac{n_h m_{hi}}{N_h M_{hi}} \times \frac{1}{M_{hi}}$$

여기서 N_h 은 h 층의 모집단 조사구 총수를 나타내며, n_h 은 h 층의 표본조사구 총수, M_{hi} 는 h 층의 i 번째 조사구에 있는 총 가구수, m_{hi} 은 h 층의 표본가구수를 나타낸다.

따라서 비례배분 하에서 표본으로 추출된 가구의 가중치(w_{hS} : sampling weight)는 다음과 같다.

$$w_{hS} = \frac{1}{p_{hij}} = \frac{N_h}{n_h m_h} \times M_{hi}$$

한편 표본추출틀로 2005년 인구총조사 90%조사구 자료를 이용하였기 때문에 최종 가중치에 (10/9)의 승수 인자를 곱하여 실제 가중치로 적용한다.

따라서 2009년 표본설계에서는 전국 또는 층별로 아동 수 총계에 대한 추정치 계산이 가능토록 하였고, 추정량과 추정량의 표준오차공식을 제시하였다.

보육실태조사에서 각종 통계치는 사후층화 조정을 통해서 얻어진 가중치를 이용하여 계산된 것으로 추정량의 분산 계산이 복잡하다. 최근에는 복합표본조사 데이터를 전문적으로 분석할 수 있는 다양한 통계소프트웨어들이 개발되어 널리 활용되고 있다. 널리 사용되고 있는 대표적인 복합표본조사 분석용 통계소프트웨어는 SAS 8.0 판 이상, SUDAAN, WestVarPC, Stata 등이다. 보육실태조사 자료에 대한 고급분석을 위해서는 복합표본조사 분석용 통계소프트웨어를 활용하는 것이 필요하다.

실제 조사를 마치고 얻은 자료들에는 무응답(예를 들면 부재, 응답거절 등)으로 인해 결측치가 생긴다. 또한 조사 결과의 특성치별(예를 들면, 성별, 연령별 등) 비율이 모집단의 비율을 따르지 않는 경우 사후층화 방법에 의해 이들을 가중치 조정에 반영해 주어야 추정의 정확성을 높일 수 있다.

제2절 특성별 추정

1. 전국 아동 수 총계 추정

복합표본조사 데이터를 분석할 때 가중치를 무시하고 분석하면 모수 추정에 심각한 편향(bias)이 발생할 수 있고, 추정량의 분산이 과소평가되어 문제가 된다. 따라서 2009년 보육실태조사에서 모집단의 특성치에 대한 추정은 가중치를 이용한 통계치를 이용해야 하고, 만약 단순총계를 사용하면 추정치에 편향이 발생할 수 있다.

2009년 보육실태조사에서 모집단의 특성치에 대한 추정은 가중치를 이용하는데, 이 때 사용될 기호들은 다음과 같다.

$h = 1, 2, \dots, L$: 층을 나타내는 첨자

$i = 1, 2, \dots, n_h$: h 층 내의 집락을 나타내는 첨자

$j = 1, 2, \dots, m_{hi}$: h 층의 i 번째 집락 내의 가구를 나타내는 첨자

n_h : h 층의 1단계 표본 집락수(표본조사구 수)

m_{hi} : h 층의 i 번째 표본 집락의 표본 가구수

$n = \sum_{h=1}^L n_h$: 표본 집락 총수(표본 조사구 총수)

$m = \sum_{h=1}^L \sum_{i=1}^{n_h} m_{hi}$: 총 표본 가구수

w_{hij} : h 층의 i 번째 집락 내의 j 번째 관찰치의 가중치

y_{hij} : h 층의 i 번째 집락 내의 j 번째 관찰치

f_h : h 층의 조사구 추출률

2009년 보육실태조사에서 전국 아동 수 총계($\hat{Y}$)는 다음과 같다.

$$\hat{Y} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$$

그리고 전국 아동 수 총계의 분산추정량은 다음과 같다.

$$var(\hat{Y}) = \sum_{h=1}^L \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi.} - \overline{y_{h..}})^2$$

여기서, $y_{hi.} = \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$, $\overline{y_{h..}} = \left(\sum_{i=1}^{n_h} y_{hi.} \right) / n_h$ 이다.

또한, 전국 아동 수 총계의 상대표준오차는 다음의 식을 통해서 계산한다.

$$\widehat{CV}(\hat{Y}) = \frac{\sqrt{\text{var}(\hat{Y})}}{\hat{Y}} \times 100 (\%)$$

2. 각 특성별 아동 수 총계 추정

2009년 보육실태조사에서 각 특성별(권역별) 아동 수 총계($\hat{Y}_h$)는 다음과 같다.

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$$

그리고 각 층별 아동 수 총계의 분산추정량은 다음과 같다.

$$\text{var}(\hat{Y}_h) = \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi.} - \overline{y_{h..}})^2$$

또한, 각 층별 아동 수 총계의 상대표준오차는 다음의 식을 통해서 계산한다.

$$\widehat{CV}(\hat{Y}_h) = \frac{\sqrt{\text{var}(\hat{Y}_h)}}{\hat{Y}_h} \times 100 (\%)$$

3. 아동의 보육수요율 추정

전국 아동의 보육수요율(R)의 추정치는 다음과 같다.

$$\hat{R} = \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} x_{hij}}$$

여기서 y_{hij} 는 h 층의 i 번째 조사구의 j 번째 가구의 보육서비스 수요 아동수이고, x_{hij} 는 h 층의 i 번째 조사구의 j 번째 가구의 아동수를 나타낸다.

그리고 전국 아동 보육수요율의 잭나이프 추정량은 다음과 같다.

$$var_J(\hat{R}) = \frac{1}{k(k-1)} \sum_{\alpha=1}^k (\hat{R}_\alpha - \hat{\bar{R}})^2$$

여기서 $\hat{R}_\alpha$ 는 α 번째를 제외한 나머지 값으로 추정한 비 추정값이며, $\hat{\bar{R}}$ 은 k 개의 그룹에 대한 비추정량의 평균이다.

또한, 전국 아동 보육수요율의 상대표준오차는 다음의 식을 통해서 계산한다.

$$\widehat{CV}(\hat{R}) = \frac{\sqrt{var_J(\hat{R})}}{\hat{R}} \times 100 (\%)$$

한편 각 층별 아동보육수요율(R_h)의 추정치는 다음의 결합비 추정량으로 정의할 수 있다.

$$\hat{R}_h = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} x_{hij}}$$

그리고 각 층별 아동보육 수요율의 잭나이프 분산추정량은 다음과 같다.

$$var_J(\hat{R}_h) = \frac{1}{k_h(k_h-1)} \sum_{\alpha=1}^{k_h} (\hat{R}_{\alpha h} - \hat{\bar{R}}_h)^2$$

여기서 k_h 는 h 층에서 k 개의 그룹을 의미하며, $\hat{R}_{\alpha h}$ 는 h 층의 α 번째 추정치를 제외한 비추정치이고, $\hat{\bar{R}}_h$ 는 h 층에서 k 개의 그룹에 대한 비추정의 평균이다.

또한, 각 층별 아동 보육 수요율의 상대표준오차는 다음의 식을 통해서 계산한다.

$$\widehat{CV}(\hat{R}_h) = \frac{\sqrt{var_J(\hat{R}_h)}}{\hat{R}_h} \times 100 (\%)$$

제5장 가중치 조정방법

표본조사의 주된 목적은 모집단으로부터 확률추출된 표본자료를 통해 미지의 모집단의 특성을 추측(inference)하는데 있다. 이때 표본을 추출하는 방법은 다양한 확률표본추출 방법을 적용할 수 있을 것이다. 이와 같이 표본으로 추출된 추출단위들은 모집단에 있는 표본으로 추출되지 않은 단위들을 잘 대표해야 한다. 즉, 표본자료는 모집단에 대한 대표성을 확보해야 하며, 이때 고려할 수 있는 것이 표본자료에 대한 가중(weighting)이다. 즉, 모집단으로부터 크기 n 의 확률표본을 단순임의 추출할 경우 모집단 총합에 대한 추정량 $\hat{Y}$ 는 $N/n \sum y_i$ 로서 i 번째 단위에 대한 승수가 N/n 이 되며 이를 i 번째 단위에 대한 가중치라 한다.

결과적으로 가중의 주된 목적은 표본자료의 모집단에 대한 대표성 확보라 할 수 있다.

이러한 가중은 첫째, 분산과 비용을 감소시킬 수 있으며, 둘째 영역별로 서로 다른 추출률을 할당하여 작은 영역의 추출률을 높임으로서 전체적인 비용을 감소시킬 수 있다. 셋째, 중복, 비포괄성 등의 불완전 추출률 문제를 해결할 수 있다. 넷째, 조사에서 발생하는 무응답문제를 적절히 해결할 수 있다. 그런데, 이와 같은 가중 절차에서 추출단위에 항상 상수의 가중치를 고려할 수 있는데, 모집단 단위가 표본으로 뽑힐 확률이 0이 아닌 값을 갖는 경우로서 이와 같은 가중치를 “자체가중(self weighting)”이라 한다. 이러한 자체가중은 단순임의추출, 층화추출, 확률비례추출, 다단계 추출 등에서 적용할 수 있다. 자체가중방법은 첫째, 조사의 복잡성을 감소시키며, 둘째 잘못된 가중치의 문제를 해결할 수 있고, 셋째, 동일한 표본이 다양한 목적과 다른 조사에서 사용될 수 있는 유연성과 편리성을 가지며, 넷째, 다양한 표본설계에 로버스트(robust)하며, 다섯째 일반적인 가중절차에 비해 이해하기가 쉽다는 장점을 가진다.

일반적으로 조사단위에 대해 가중의 절차는 (1) 추출가중치 또는 기본가중치 (2) 무응답가중치 (3) 사후층화가중치의 단계로 이루어진다.

제1절 기본가중치

기본가중치(base weight)는 표본추출 설계로부터 직접적으로 얻어지는 값이다. 임의의 모집단으로부터 적절한 크기의 표본을 추출할 때 단위가 표본에 포함될 확률의 역수로 기본가중치를 계산할 수 있다. 이때 포함확률은 기지(known)의 값이다. 다음은 표본 각 추출설계에 따른 기본가중치를 산정한 것이다.

〈표 4-1〉 추출설계와 가중치

추출설계	가중치
단순임의추출(비복원)	$w_i = N/n$
단순계통추출(추출간격이 r 인 경우)	$w_i = r$
층화임의추출(전체 L 개의 층)	$w_{hi} = N_h/n_h$
계통 추출(불균등확률)	$w_i = M/(nM_i)$
2단계 추출 $p_i^{(1)}$: i 번째 PSU의 추출확률 $p_{ij}^{(2)}$: i 번째 PSU에서 j 번째 SSU의 조건부추출확률	$w_{ij} = 1/(p_i^{(1)} p_{ij}^{(2)})$

제2절 무응답 조정가중치

무응답 조정가중치(nonresponse weighting)의 중요한 역할은 조사로부터 발생한 무응답으로 인한 무응답 편향을 제거하기 위한 것이다. 무응답 편향은 무응답자들이 응답자들과 차이가 있을 때 발생하며, 모집단에 대해 매우 높은 비율로 추정치에 영

향을 줄때 발생한다. 무응답 조정가중치를 계산하는 방법으로는 이용 가능한 정보의 근원에 따라 “표본에 기초한 무응답조정 방법”과 “외부정보를 이용한 무응답 조정방법”이 있다.

1) 표본에 기초한 무응답 조정방법

이용 가능한 정보가 표본으로 한정되며, 전체 모집단에 대한 정보는 알 수 없고, 무응답 단위들의 기본가중치를 표본응답자들에게 배정하여 응답단위들에 대해 조정된 가중치의 합이 전체 표본단위들에 대한 기본가중치의 합이 된다.

$$F_c = \frac{\sum_{i=1}^{n_1} w_i MOS_i + \sum_{i=1}^{n_2} w_i MOS_i}{\sum_{i=1}^{n_1} w_i MOS_i}$$

여기서 n_1 은 응답자들의 수이고, n_2 는 무응답자들의 수이다. w_i 는 기본가중치이고 MOS_i 는 표본추출설계에 의해 결정되는 적절한 크기척도를 나타낸다.

그러면 조정값 F_c 를 표본응답자들의 기본가중치를 곱하여 각 단위의 가중치로 고려하면 무응답 조정 가중값 $w_i^{(a)}$ 를 계산할 수 있다.

$$w_i^{(a)} = \begin{cases} F_c w_i, & \text{그룹1} \\ 0, & \text{그룹2} \\ w_i, & \text{그룹3} \end{cases}$$

이때 그룹1은 응답그룹을 그룹2는 무응답그룹을 그룹3은 부적절한 표본그룹을 나타낸다.

2. 외부정보를 이용한 무응답 조정방법

표본에 기초한 무응답 조정 가중치를 계산한 후 외부자료를 이용하여 사후층화, raking, 또는 calibration 등의 방법으로 이 가중치를 조정한다. 외부자료를 이용한 경우 조정인자 F_c 가 다음과 같이 수정된다.

$$F_c^* = \frac{MOS_c}{\sum_{i=1}^{n_1} w_i MOS_i}$$

여기서 MOS_c 는 계급 c 에 대해 외부데이터로부터 얻는 크기척도이다.

이때 유일한 외부정보는 표본 추출틀로부터 구할 수 있으며 이를 가중계급을 구성하는데 이용한다. 가중계급을 구성하기 위해 이용 가능한 방법으로는 경험적인 방법, 분류소프트웨어(CHAD), 성향점수모형, 다중가법회귀나무(MART) 등이 있다.

제3절 사후층화 가중치

1. 사후층화 조정

사후층화(post-stratification) 조정은 추출틀의 불완전으로 인한 포괄성의 차이, 표본의 불균형 또는 비대표성, 무응답에 의한 차이 등을 조정하기 위해 광범위하게 이용되는 방법이다. 즉, 표본응답자들의 가중치를 조정함으로써 가중된 표본분포가 기지의 모집단분포와 같아지도록 하는 방법이다. 이러한 사후층화 조정을 실시하는 주된 이유는 첫째, 추정치의 정도를 개선할 수 있으며, 둘째, 추정분산을 줄이기 위해 층화와 비추정을 사용할 수 있고, 셋째, 부차모집단간의 포괄성과 무응답에 따른 추정치의 편향을 감소시키며, 넷째, 모집단의 다양한 그룹에 대해 추정치의 일치성을 보장한다.

사후층화를 위한 초기 셀 구성의 원칙은 다음과 같다.

- ① 모집단 수는 부차 데이터로부터 이용 가능해야 하며 이들의 특성 값들은 관심변수와 양의 상관관계를 가져야 한다.
- ② 모집단 수는 정확히 목표모집단 수가 되어야 한다.
- ③ 모집단 수는 가중된 조사추정치보다 더욱 신뢰할 수 있어야 한다.
- ④ 하나의 표본응답자는 하나의 셀로 분류될 수 있어야 한다.
- ⑤ 표본에 대한 사후층의 정의와 기지의 모집단은 서로 일치해야 한다.
- ⑥ 사후층 셀들은 충분히 커야 한다.

2. 래킹비 조정

래킹비조정(raking ratio adjustment)은 전수조사 자료와 표본조사 자료간의 일치성을 확보하기 위해 1940년에 미국에서 Deming과 Stephan에 의해 처음 제안된 방법으로 2차원 분류표상의 각 셀 값을 반복적으로 조정해 가는 방법이다. 래킹비 조정 절차는 기본 가중치를 하나의 주변분포를 이용하여 조정한 후, 두 번째 주변분포를 재차 이용하여 가중치를 조정한다. 이러한 과정을 특정한 수렴조건을 만족할 때까지 반복적으로 수행한다.

래킹비 조정 절차의 단점은 첫째, 반복적으로 셀 값을 조정해가는 과정에서 수렴성을 보장할 수 없다는 점이다. 즉, 래킹해야 할 차원을 잘못 선택할 경우 래킹 과정이 수렴하지 않을 수 있다. 둘째, MSE 계산이 매우 어렵다는 점이다. 왜냐하면 래킹 추정량의 구조가 복잡하기 때문에 직접적으로 MSE를 계산하기 어렵다는 문제가 있다. 셋째, 무응답과 같은 상황에 바로 적용하기란 쉽지 않다.

3. 보정

보정(calibration)은 추출설계에 따른 기본가중치 또는 추출가중치와 보정된 새로운 가중치 간의 차이를 나타내는 일종의 거리함수(distance function)를 주어진 조건을 만

축하도록 최소로 하는 가중값을 구한 방법이다. 이때 제한조건을 나타내는 식을 보정방정식(calibration equation)이라 하며 다음과 같이 정의한다.

$$\sum_s w_i x_i = \sum_U x_i = \tau_x$$

여기서 x_i 는 관심변수 y_i 와 강한 상관이 있는 보조변수이다.

보정방정식의 조건하에서 특정한 거리함수를 최소로 하는 새로운 가중치 w_i 를 구하면 된다.

즉, 선형 거리함수 $\sum G(w, d) = \sum (w_i - d_i)^2 / 2d_i$ 를 최소로 하는 원래의 추출 가중치 d_i 의 새로운 가중치 w_i 를 구하면 $w_i = d_i (1 + \sum d_i x_i) (\sum x_i x_i')^{-1} (\tau_x - \hat{\tau}_x)$ 이며 이를 총합추정량의 식에 대입하면 다음과 같은 일반화 회귀추정량(generalized regression estimator : GREG)이 된다.

$$\widehat{\tau_{yGREG}} = \widehat{\tau_{yHT}} + \hat{B}' (\tau_x - \widehat{\tau_{xHT}})$$

이와 같은 회귀추정량의 장점은 연속형 보조정보를 사용할 수 있고, 유연성을 가지며, 분산을 정의할 수 있다는 점이다.

제4절 가중의 효과

표본의 추출률과 조사에서 발생하는 무응답, 그리고 프레임의 불완전성으로 인한 비포괄성 문제 등을 해결하기 위해서는 반드시 가중치를 고려해야 한다. 만일 가중을 하지 않으면, 추정치의 편향이 증가하게 되며 따라서 추정치의 왜곡을 피할 수 없게 된다. 그러나 가중치를 고려하게 되면 추정치의 편향은 감소하지만 가중의 효과로 인해 분산이 증가하게 된다. 이러한 이유 중의 하나는 각 단위에 부과되는 가중치의 변동이 매우 클 경우에 발생한다.

모집단 평균을 추정함에 있어 분산의 증가분에 기여한 가중치의 효과는 다음과 같은 인자에 의해 측정된다.

$$L = n \times \frac{\sum_h n_h w_h^2}{(\sum_h n_h w_h)^2}$$

여기서 $n = \sum_h n_h$ 로서 실제 표본크기를 나타내며, w_h 는 최종가중치, n_h 는 h 층의 실제 표본크기이다.

이 식은 다음과 같이 가중치의 변동계수(CV)의 식으로 재표현이 가능하다.

$$L = n \times \frac{\sum_j w_j^2}{(\sum_j w_j)^2} = 1 + CV^2(w_j)$$

이때 $CV^2(w_j) = \frac{n}{(\sum_j w_j)^2} \left(\sum_j w_j^2 - \frac{1}{n} (\sum_j w_j)^2 \right)$ 이다.

제5절 가중치의 절단

1차적으로 최종적인 가중치가 계산되어 불완전성을 보상하기 위해 조정된 다음, 조정된 가중치의 분포를 파악하는 것이 바람직하다. 극히 작은 표본에 의해 극단적으로 큰 가중치는 추정치의 분산을 증가시키는 요인이 되기 때문이다. 그러므로 가중치의 변동을 고려하여 최대값 수준에서 극단가중치를 절단하는 것이 필요하다. 가중치의 절단(trimming of weights)은 대체로 무응답에 대한 조정 후에 수행하는 것이 일반적이다.

가중치의 절단 작업은 분산을 축소시키는 효과가 있지만, 다른 한편으로는 추정치

의 편향을 야기하게 된다. 따라서 매우 큰 가중치에 대해 절단 작업을 수행하면, 분산의 크기는 줄일 수 있지만, 편향의 크기가 상대적으로 증가하게 되는 문제가 있다. 따라서 가중치의 절단은 절단을 함으로서 발생하는 분산의 감소분 보다 총 MSE 상의 영향이 적도록 가중치를 절단하는 것이 바람직하다.

층화 추출설계에 대해, 가중치 절단과정은 각 층 내에서 수행되는 것이 이상적이다. 먼저 원 가중치들에 대한 상한(upper bound)를 정의하고, 전체 가중치들에 대해 절단된 가중치들의 합이 원가중치들의 합과 같아지도록 조정한다. w_{hi} 를 h 층의 i 번째 단위에 대한 최종 가중치라 하고, w_{hB} 를 h 층에 대해 정해진 가중치들의 상한이라 하자. 그러면, h 층의 i 번째 표본단위에 대한 절단 가중치는 다음과 같다.

$$w_{i(T)} = \begin{cases} w_{hi}, & w_{hi} \leq w_{hB} \\ w_{hB}, & w_{hi} \geq w_{hB} \end{cases}$$

그러면 전체 표본에 대해 절단된 가중치들의 합이 원 가중치의 합과 같아지도록 조정되어야 한다. 층 내에서 상수인 가중치를 가정하고, F_T 를 원 가중치들의 합과 절단된 가중치들의 합의 비라 하자.

$$F_T = \frac{\sum_h n_h w_h}{\sum_h n_h w_{h(T)}}$$

만일 h 층에 대해 조정된 절단가중치를 $w_{h(T)}^*$ 이라 하면 이는 다음과 같다.

$$w_{h(T)}^* = F_T \times w_{h(T)}$$

이때 $w_{h(T)}^*$ 는 $\sum_h n_h w_{h(T)}^* = \sum_h n_h w_h$ 를 만족한다.

제6절 가중치 조정시 고려 사항

가중치 조정시에 고려해야 할 사항을 정리해 보면 다음과 같다.

□ 표본설계시 자체가중설계를 고려하는 것이 바람직하다.

자체가중설계(self-weighting design)의 장점은 이해가 쉽고 다른 추출설계에 직접적으로 적용이 가능하다는 점이다. 따라서 표본추출 설계시 가능하면 자체가중설계를 고려함으로써 표본설계의 유연성을 확보할 수 있다.

□ 자체가중 설계가 어렵다면 가중치 조정단계에 따라 가중치를 부여해야 한다.

만일 자체가중 설계가 표본설계과정에서 고려되지 않았거나 고려하기 어려운 경우라면, 기본가중치와 무응답가중치 그리고 사후가중치 조정단계에 따라 가중치를 계산하여 추정치 산정에 반영해야 할 것이다. 또한 가중치 조정단계에서 이용 가능한 정보를 적극 활용하는 것도 고려해야 한다.

□ 가중치 조정과정에서 이용 가능한 보조정보의 활용을 검토해야 한다.

이용 가능한 정보는 표본과 표본 외부에 모두 존재할 수 있으며, 만일 외부정보를 활용할 수 있다면 이를 적극적으로 활용하는 것이 바람직하다. 사후층화 조정(post-stratified adjustment)시에 이러한 외부정보를 사용함으로써 표본과 모집단간의 일치성을 보장할 수 있을 것이다.

□ 가중치를 적용했다면 가중치 효과를 반드시 계산해야 한다.

통상적으로 가중치를 사용하지 않으면, 추정량의 편향이 발생하게 되는 반면에 가중치를 사용하면 추정량의 편향은 줄일 수 있으나, 분산이 증가하는 문제가 있다. 따라서 무조건 가중치를 적용할 것이 아니라 먼저 가중효과를 분산팽창인자(variance inflation factor: VIF)를 이용하여 계산해서 적절한 가중치를 적용해야 할 것이다.

- 가중치의 변동을 고려하여 극단 가중치는 추정치에 큰 영향을 줌으로 객관적인 방법으로 절단하여 적용하는 것도 고려할 수 있다.

극단 가중치는 분산을 크게 함으로서 추정치의 정도를 떨어뜨리는 역할을 한다. 따라서 적절한 기준 하에서 극단가중치를 조정할 필요가 있다. 이는 총 평균제곱오차(MSE) 차원에서 다루는 것이 합당한데, 왜냐하면 가중치를 조정하여 분산감소의 효과를 볼 수 있는 반면 추정치의 편향이 발생하기 때문이다.

제6장 조사 및 비표본오차의 관리

2004년도 표본설계는 2000년 인구주택총조사에서 표본조사가 실시된 조사구를 모집단조사구로 사용한다. 이들은 전체 모집단의 10%에 해당된다. 모집단조사구에서 표본조사구를 추출하고 추출된 표본조사구에 있는 가구 중에서 일부를 표본 추출한다. 추출된 표본 가구의 가구원 중에서 영유아 및 미취학 아동이 있는 가구를 대상으로 조사를 실시하였다. 조사목적이 만 6세 이하의 영유아 및 미취학 아동의 보육 및 교육실태 및 욕구를 파악하기 위한 조사이므로 영유아 및 미취학 아동이 없는 가구에 대한 조사는 기본사항만을 조사하였다.

2009년도의 전국 보육실태조사의 경우 2005년도의 자료를 사용하는 관계로 그동안 모집단의 변화가 있으리라 예상되므로 표본으로 선정된 조사구나 표본가구에 일부 변동이 있을 것이다. 그러므로 이러한 변동을 파악해서 조사에 반영해 주어야 하며, 또한 조사 과정에서 발생하는 오차를 잘 관리 해주어야 조사의 정확성을 높일 수 있다. 최근 들어 조사환경이 열악해져서 조사의 어려움이 가중되어 비표본오차의 비중이 커지고 있다. 따라서 이들 비표본오차의 발생 원인을 살펴서 비표본오차의 발생을 최소화 한다.

제1절 조사관리

2009년도 보육실태조사 표본설계에는 2005년에 실시한 인구주택총조사에서 전수조사가 실시된 조사구를 모집단조사구로 사용하므로 전체 조사구의 90%가 된다. 2009년 조사에서는 이와 같이 2005년도 모집단 자료의 변화가 있을 것으로 예상되므로 표본으로 선정된 조사구와 표본조사구내의 가구에 대한 확인 작업이 필요하다. 표본으로 선정된 조사구가 없어지거나 조사에 사용될 수 없는 경우에는 예비 표본

조사구를 사용한다. 또한 표본조사구내의 가구를 확인해서 가구목록을 새로이 작성해야 한다. 이와 유사한 2007년도 “차상위 실태조사”의 경우 부재 및 재개발등의 사유로 조사가 불가능한 경우가 발생하였고, 특히 특정 조사구의 경우 조사 거절도 많아서 실제 접촉가구의 성공률이 낮았으므로 이번 조사에서는 동일한 오류를 범하지 않도록 각별한 주의가 필요하다. 주요 국가들에서 가계조사 무응답률 현황은 <표 5-1>과 같다.(참고; 최봉호, 2006) 아직까지 우리나라에서 가계조사의 무응답률은 다른 국가들에 비해 낮은 편이나 점차적으로 높아져서 조사에 어려움이 커지고 있다.

〈표 5-1〉 주요 국가의 가계조사 무응답률

(단위: %)

	한국	미국	캐나다	호주	영국	프랑스	아일랜드	폴란드	헝가리
연도	2005	2001	2004	1998/99	2000/01	1994/95	1999/00	2000	2000
무응답률	16.5	25.1	30.8	22.6	41.2	27.0	45.0	49.2	39.0

또한 통계청에서 2006년 5월 기준으로 조사원 중 일부(472명)를 대상으로 조사환경에 대한 설문조사를 실시한 자료가 <표 5-1>에 있다. <표 5-2>에서 조사원들이 경상적인 표본조사를 위해 조사대상처를 방문할 때, 절반 정도는 방문 당일 조사를 완료하고, 나머지는 조사 불응 또는 가구원 부재로 인해 조사를 하지 못하고 있음을 보여 주고 있다. 반면, 농어가 경제조사 및 사업체대상 조사의 경우는 면접완료율이 가계조사의 경우보다 다소 높은 68% 수준을 나타내고 있다(참고; 최봉호, 2006).

〈표 5-2〉 통계청 경상표본조사에서의 당일 방문 면접완료율

(단위: %)

	계	면접 완료	조사 불응	가구원 부재
경제활동 인구조사	100.0	50.6	3.3	46.1
가계조사	100.0	46.9	13.1	40.0
농/어가 경제조사	100.0	68.0	-	32.0
사업체 대상 조사	100.0	68.3	13.0	24.6

2009년도 『전국보육실태조사』는 만 6세 이하 미취학 아동의 보육실태를 파악하기 때문에 통상적으로 대상가구를 지정하지 않는 통계청의 가계조사보다 영유아 가구의 출현율을 일정수준 확보해야 하기 때문에 어려움이 더할 것으로 예상된다. 따라서 실사에서 만전의 준비가 요구된다. 일차적으로 응답률을 높이는 조치가 필요하고, 무응답이 발생한 경우 이를 보완해 주는 추가적인 작업이 있어야 한다. 표본이 모집단을 충분히 대표하지 못하는 경우 가중치 조정을 통해서 추정의 정확성을 높인다. 다음 절에서는 일반적으로 널리 사용되고 있는 가중치 조정방법을 소개한다. 본 조사를 마무리 한 후 적절한 가중치 조정방법을 사용하여 추정의 정도를 높이도록 한다.

제2절 비표본오차 관리

통계조사에서 발생하는 오차는 크게 표본오차와 비표본오차로 구분된다. 표본오차는 모집단으로부터 일부를 표본으로 추출함으로써 생기는 오차이다. 반면에 비표본오차는 표본오차 이외의 모든 오차로 발생 원인도 다양할 뿐만 아니라 측정하기도 어려워서 비표본오차가 발생하지 않도록 사전에 조치를 취하는 것이 무엇보다도 중요하다.

조사 설계와 과정에서 발생하는 오차의 유형과 발생 원인들을 꼼꼼히 살펴서 비표본오차의 발생을 줄이도록 한다.(참고; 한국조사연구학회, 2007)

1. 포함오차(coverage error)

표본추출틀은 표본을 추출하는 데 사용되는 목표모집단 구성 요소들의 목록이다. 포함오차는 표본추출틀이 목표모집단과 일치하지 않는 경우에 발생한다. 포함오차가 발생하는 경우는 다음과 같이 세 가지로 구분할 수 있다.

- 첫째, 목표모집단의 일부 단위가 표본추출틀에 포함되지 않는 경우로, 이를 과

소포함(undercoverage) 또는 미포함(noncoverage)이라 한다.

- 둘째, 표본추출틀에 있는 일부 단위가 목표모집단의 구성원이 아닌 경우, 이들은 부적격 또는 조사 범위를 벗어나는 것이므로 표본추출 전에 확인하고 이들을 표본추출틀에서 제거한다.
- 셋째, 표본추출틀에서 두 개 이상의 단위가 목표모집단에서 동일한 단위와 대응이 될 때, 이를 과대포함(overcoverage)이라 한다.

본 조사에 사용되는 표본추출틀은 2005년도에 실시한 인구주택총조사에서 전수조사가 실시된 90%의 자료를 사용한다. 따라서 2009년도 조사에서는 일부 포함오차가 발생할 것으로 사료된다. 그리고 본 조사는 매년 실시되는 조사로 해가 거듭될수록 포함오차가 증가할 것으로 예상된다. 이런 경우 조사 전에 가구의 변동을 파악해서 표본이 모집단을 충분히 반영할 수 있도록 새로운 자료를 보완하거나 조정해서 포함오차가 증가하지 않도록 한다.

2. 무응답오차(nonresponse error)

무응답에는 단위무응답(unit nonresponse)과 항목무응답(item nonresponse)이 있다. 표본으로 추출된 조사단위(가구 또는 가구원)가 설문에 응답을 하지 않은 경우를 단위무응답(unit nonresponse)이라 하고, 조사에 응한 응답자가 일부 조사항목에 응답을 하지 않은 경우를 항목무응답(item nonresponse)이라 한다. 5장 1절에서 언급한 바와 같이 최근 들어서 조사 환경의 변화로 표본으로 선정된 조사단위가 조사에 응하지 않거나 응답을 거부하는 경우가 자주 발생하므로 이에 대한 대비를 해야 한다. 무응답이 발생하는 원인을 사전에 제거하는 것이 중요한데, 일단 무응답이 발생한 경우에는 5장 2절을 참고하여 적절한 조치를 취해서 오차를 최소화 시키도록 한다. 더불어서 표본으로 선정된 가구를 대상으로 안내문을 보내거나 적절한 보상을 하여 최대한 응답률을 높이는 노력이 필요하다.

3. 측정오차(measurement error)

측정오차는 관측값과 참값과의 차이로 비표본오차의 발생 요인 중에서 가장 많은 연구가 진행되어 왔다. 측정오차를 정의하는 방법은 연구자의 참값에 대한 입장에 따라 차이가 난다. 참값이 통계조사 상황에 관계없이 독립적으로 존재한다는 것과 참값이 통계조사 상황과 연관시킬 때만 의미가 있다는 각기 다른 관점에서 접근하는 방법이다.

측정오차를 유발시키는 주요 원인은 응답자, 조사원, 그리고 조사표이다. 이들 원인들이 자료 수집 전체에서 발생하고 각 원인들이 측정과정에서 오차를 생기게 한다. 예를 들어, 응답자가 고의적으로 또는 무심코 틀린 정보를 줄 수 있다. 또한 조사원들이 자료를 왜곡시키거나, 응답에 부적절한 영향을 주거나, 응답을 잘못 기록하는 경우에도 측정오차가 발생한다. 조사표가 잘못 설계되어 애매한 질문을 하게 되거나 혼동되는 지시, 오해하기 쉬운 용어들이 조사표에 있으면 측정오차가 생긴다. 측정오차는 계량화하기가 어렵고 비용이 많이 들기 때문에 특별한 연구가 필요하다. 재조사계획, 기록점검, 코딩, 인지검사(cognitive testing), 랜덤실험 등과 같은 방법들이 측정오차를 계량화하는데 사용된다.

[그림 6-1] 조사연구과정에서 발생하는 오차



4. 처리오차(processing error)

처리오차는 조사 자료가 수집된 후에 발생하는 오차이다. 자료의 수집에서 최종 조사결과 발표에 이르는 각 처리 단계에서 오류가 발생할 수 있다. 자료 에디팅, 자료입력, 코딩, 가중값, 조사표 작성 등을 포함하는 단계에서도 발생한다. 이들 오류들은 단순한 기록, 복사 또는 전송 오류에서부터 전문성이 부족한 편집(edit)이나 잘못된 대체 모형에 의한 복잡한 오류까지 그 범위가 다양하다.

코드화될 서술형 항목에서는 코딩오차가 또 다른 종류의 처리오차이다. 응답에 코드 번호를 부여하는 경우 코드부여 체계 자체가 불명확하여 오류가 발생할 수도 있다. 가중값을 계산하거나 여러 종류의 결과표를 작성하는 소프트웨어에 있는 오류도 결과에 영향을 주므로 이들도 처리오차의 범주에 들어간다.

5. 조정보차(adjustment error)

[그림 1]의 마지막 단계는 조사 결과로 얻어진 표본추정치를 개선하려는 노력과 관계된다. 조정을 위해서는 목표모집단(또는 조사모집단)이나 응답률과 같은 정보들을 사용한다. 예를 들어 무응답에 대한 조정을 하고자 할 때 표본자료가 대표성이 떨어진 경우, 즉 응답률이 낮은 경우에는 그 자료에 가중치를 더 크게 해주고 응답률이 높은 경우에는 가중치를 낮게 해준다. 모집단 평균과 가중치를 주어 조정한 평균과의 오차는 표본들이 달라지면 변하게 된다. 즉, 조정은 일반적으로 추정치의 편향(bias)과 분산에 영향을 준다. 무응답 조정에 대한 상세한 것은 5장을 참고하면 된다.

제7장 제언

우리나라의 아동 보육 정책 결정에 기초 자료가 되는 『전국보육실태 조사』의 신뢰성 확보와 조사 자료의 효율적인 활용을 위해서 다음과 같은 몇 가지 사항에 대한 연구검토를 제안한다.

☐ 통계전문가에 의한 조사계획과 조사 자료의 관리

『전국보육실태 조사』에서 데이터의 질을 향상시키기 위해서는 조사계획 단계부터 통계전문가의 활용이 필요하다. 또한 정확하고 다양한 정보를 습득하기 위해서는 통계전문가에 의해 조사 자료에 대한 관리가 우선적으로 이루어져야 한다.

☐ 모집단의 변화를 적시에 반영해 주는 주기적 표본설계

표본이 모집단을 잘 반영하고, 또한 변화하는 모집단의 특성이 표본에 반영되어 조사결과의 정확성과 신뢰성을 높이기 위해서는 모집단의 변화를 적시에 반영해 주는 주기적인 표본설계가 필요하다. 이를 통해 체계적이고 지속적인 표본설계를 통해 질 높은 통계 생산이 가능하다.

☐ 표본조사구 유지 및 확대

표본의 규모가 줄어드는 관계로 표본의 대표성이 떨어지게 되며, 조사의 정확성도 문제가 된다. 이를 방지하기 위해서 표본가구수는 줄지만 표본조사구수를 유지하거나 늘리는 방법을 사용한다. 이는 줄어드는 표본규모로 인한 대표성의 문제와 조사 자료의 질 저하를 방지하기 위함이다.

☐ 조사 자료에 대한 다양한 통계분석

조사 자료의 처리와 분석에 있어서도 통계전문가들을 활용하여 정확하고 다양한

통계생산이 가능토록 해야 한다. 과거에는 조사결과에 대한 단순한 집계만을 하여 통계분석 과정에서 편향이 발생할 수 있었다. 따라서 단순집계가 아닌 가중치를 고려한 추정치들이 산출될 수 있도록 조사 자료에 대한 분석이 이루어져야 한다. 또한 조사 자료에 대한 다양한 통계적 분석과 추정치와 추정치의 표준오차들이 계산되어져야 한다.

새로운 표본설계는 기존의 2004년 조사와는 달리 2009년 조사에서는 보육실태만을 조사하도록 하고 있기 때문에 조사 자료의 시계열 유지가 곤란한 측면이 있다. 따라서 향후 조사에서는 보다 엄밀한 조사 기획 및 정책 효과성을 분석하여 조사자료의 내용적인 측면에서 시계열을 유지할 수 검토할 필요가 있다.

또한 비용이나 조사환경 등 여러 가지 사정으로 조사구를 늘이거나 표본 수 확대가 어려울 것으로 생각된다. 이러한 문제로 인해 지역통계를 생산하지 못하는 어려움이 있다. 사정이 허락된다면, 향후 5년에 한 번 실시되는 조사이므로 예산 확충을 통해 지자체들이 요구하는 지역통계가 생산될 수 있도록 고려해야 할 것이다. 더불어 보육수요율에 대한 허용오차를 본 표본설계에서는 6% 수준으로 정하고 있는 데 예산 확보를 통해 표본크기를 증가시켜 허용오차를 5% 미만으로 줄이는 방안도 모색할 필요가 있다.

마지막으로 조사구 크기와 추정을 고려할 때 조사구 수를 확대하고, 조사구당 표본가구수를 축소하는 것이 바람직 할 것으로 사료된다.

참고문헌

- 김영원, 류제복, 박진우, 홍기학 공역(2006), 표본조사의 이해와 활용, 톰슨코퍼레이션코리아.
- 박홍래(2000), 통계조사론, 영지문화사.
- 한국고용정보원(2006), 2006 산업·직업별 고용구조조사 - 기초분석보고서 5차년도 자료분석.
- 한국보건사회연구원(2004), 2004년 보육 및 교육 실태조사 보고서.
- 한국보건사회연구원(2006), 표본추출 및 관리 매뉴얼.
- 한국보건사회연구원(2007), 자체통계 품질진단 매뉴얼.
- 한국조사연구학회(2004), 제7차 전국 장내 기생충감염 실태조사 표본설계.
- 한국조사연구학회(2007), 농촌통계조사 표본설계 최종보고서.
- 한국조사연구학회(2006), 산업·직업별 고용구조조사 표본설계.
- 한국통계학회 조사통계연구회(2000), 무응답 오차, 자유아카데미.
- Biemer, P. P. and Lyberg, L. E.(2003), Introduction to Survey Quality, John Wiley and Sons.
- Cochran. W. G.(1977), Sampling Techniques, John Wiley & Sons, Inc.
- Deville, J. C. , Swensson, B., and Wretmman, J. (2003), Model Assisted Survey Sampling, Springer, New York.
- Deville, J. C. , Srandal, C.E.(1992), Calibration Estimators in Survey Sampling, Journal of American Statistical Association, 87, 376-382.

- Groves, R. M., Fowler, Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R.(2004), Survey Methodology, John Wiley and Sons.
- Groves, R. M., Dillman, D. A. Eltinge, J. L., and Little, R. J. A.(2002), Survey Nonresponse, John Wiley and Sons.
- Kish, L.(1965), Survey Sampling, John Wiley & Sons, Inc.
- Lessler, J. T. and Kalsbeek, W. D.(1992), Nonsampling Error in Surveys, John Wiley & Sons, Inc.
- Thompson, S. K.(2002), Sampling, 2nd Eds. John Wiley & Sons, Inc. New York.

부록

⋮

표본 조사구 리스트



【본 조사구 목록 : 257개 조사구】

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
1	1	서울특별시	중구	신당2동	0131
2	1	서울특별시	용산구	이촌1동	022A
3	1	서울특별시	성동구	도선동	0511
4	1	서울특별시	성동구	성수1가1동	092A
5	1	서울특별시	광진구	군자동	1291
6	1	서울특별시	동대문구	제기1동	0081
7	1	서울특별시	동대문구	청량리1동	017A
8	1	서울특별시	중랑구	면목1동	0491
9	1	서울특별시	중랑구	망우2동	0051
10	1	서울특별시	성북구	동소문동	061A
11	1	서울특별시	성북구	월곡2동	0181
12	1	서울특별시	강북구	미아6.7동	081A
13	1	서울특별시	강북구	번2동	0971
14	1	서울특별시	도봉구	창4동	013A
15	1	서울특별시	도봉구	도봉1동	0731
16	1	서울특별시	노원구	공릉2동	017A
17	1	서울특별시	노원구	중계2동	035A
18	1	서울특별시	노원구	상계7동	058A
19	1	서울특별시	은평구	갈현2동	0361
20	1	서울특별시	은평구	진관외동	0531
21	1	서울특별시	서대문구	연희2동	042A
22	1	서울특별시	서대문구	북가좌2동	1031
23	1	서울특별시	마포구	연남동	1091
24	1	서울특별시	마포구	성산2동	033A
25	1	서울특별시	양천구	신정2동	054A
26	1	서울특별시	양천구	신정5동	0091
27	1	서울특별시	강서구	등촌3동	097A
28	1	서울특별시	강서구	발산2동	0161
29	1	서울특별시	강서구	방화3동	104A
30	1	서울특별시	구로구	개봉3동	1111
31	1	서울특별시	구로구	오류1동	054A
32	1	서울특별시	금천구	시흥본동	1191
33	1	서울특별시	영등포구	당산2동	122A
34	1	서울특별시	동작구	노량진1동	0331
35	1	서울특별시	동작구	사당2동	001A
36	1	서울특별시	동작구	신대방1동	0851
37	1	서울특별시	관악구	신림2동	0281
38	1	서울특별시	관악구	신림8동	014A

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
39	1	서울특별시	서초구	서초4동	1151
40	1	서울특별시	서초구	반포2동	085A
41	1	서울특별시	강남구	삼성2동	0011
42	1	서울특별시	강남구	대치1동	055A
43	1	서울특별시	강남구	일원본동	033A
44	1	서울특별시	송파구	마천2동	0981
45	1	서울특별시	송파구	송파2동	076A
46	1	서울특별시	강동구	상일동	080A
47	1	서울특별시	강동구	명일1동	0881
48	1	부산광역시	중구	대청동	0471
49	1	부산광역시	중구	광복동	004A
50	1	부산광역시	영도구	남항동	0531
51	1	부산광역시	부산진구	전포2동	112A
52	1	부산광역시	부산진구	가야1동	0381
53	1	부산광역시	동래구	복산동	084A
54	1	부산광역시	남구	대연1동	0111
55	1	부산광역시	남구	감만1동	077A
56	1	부산광역시	북구	덕천1동	0921
57	1	부산광역시	북구	덕천2동	102A
58	1	부산광역시	해운대구	중2동	062A
59	1	부산광역시	해운대구	좌제4동	134A
60	1	부산광역시	사하구	괴정4동	0171
61	1	부산광역시	사하구	다대1동	029A
62	1	부산광역시	금정구	장전2동	0011
63	1	부산광역시	연제구	거제1동	132A
64	1	부산광역시	연제구	연산9동	0681
65	1	부산광역시	사상구	모라3동	014A
66	1	부산광역시	사상구	엄궁동	1661
67	3	부산광역시	기장군	일광면	0511
68	1	대구광역시	동구	신암3동	052A
69	1	대구광역시	동구	안심1동	1281
70	1	대구광역시	남구	이천동	045A
71	1	대구광역시	남구	봉덕1동	0301
72	1	대구광역시	북구	북현2동	1411
73	1	대구광역시	북구	읍내동	024A
74	1	대구광역시	수성구	만촌3동	069A
75	1	대구광역시	수성구	파동	0711
76	1	대구광역시	수성구	고산3동	037A
77	1	대구광역시	달서구	진천동	194A
78	1	대구광역시	달서구	송현2동	0381

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
79	1	대구광역시	달서구	용산2동	085A
80	3	대구광역시	달성군	다사읍	036A
81	1	인천광역시	남구	학익1동	1281
82	1	인천광역시	남구	학익2동	012A
83	1	인천광역시	연수구	연수2동	0161
84	1	인천광역시	연수구	연수3동	119A
85	1	인천광역시	남동구	만수1동	100A
86	1	인천광역시	부평구	부평1동	0151
87	1	인천광역시	부평구	산곡2동	045A
88	1	인천광역시	부평구	삼산동	282A
89	1	인천광역시	부평구	십정1동	0651
90	1	인천광역시	계양구	작전1동	039A
91	1	인천광역시	서구	가정3동	003A
92	1	인천광역시	서구	신현원창동	1091
93	1	광주광역시	동구	학운동	043A
94	1	광주광역시	서구	화정1동	0511
95	1	광주광역시	서구	풍암동	107A
96	1	광주광역시	남구	봉선2동	087A
97	1	광주광역시	북구	풍향동	0301
98	1	광주광역시	북구	두암2동	116A
99	1	광주광역시	광산구	어룡동	059A
100	1	광주광역시	광산구	운남동	070A
101	1	대전광역시	동구	가양1동	0601
102	1	대전광역시	중구	태평1동	039A
103	1	대전광역시	서구	도마2동	0841
104	1	대전광역시	서구	둔산1동	055A
105	1	대전광역시	서구	만년동	060A
106	1	대전광역시	유성구	신성동	0231
107	1	대전광역시	유성구	노은동	084A
108	1	울산광역시	중구	우정동	043A
109	1	울산광역시	중구	다운동	0691
110	1	울산광역시	남구	무거2동	202A
111	1	울산광역시	동구	화정동	1281
112	1	울산광역시	북구	농소1동	057A
113	3	울산광역시	울주군	웅촌면	008A
114	2	경기도	수원시 장안구	파장동	027A
115	2	경기도	수원시 장안구	조원2동	060A
116	2	경기도	수원시 권선구	세류3동	0911
117	2	경기도	수원시 권선구	권선2동	100A
118	2	경기도	수원시 팔달구	고등동	0861

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
119	2	경기도	수원시 영통구	매탄4동	039A
120	2	경기도	수원시 영통구	태장동	148A
121	2	경기도	성남시 수정구	태평4동	0211
122	2	경기도	성남시 중원구	은행2동	0131
123	2	경기도	성남시 분당구	정자2동	013A
124	2	경기도	성남시 분당구	이매2동	020A
125	2	경기도	의정부시	의정부2동	0301
126	2	경기도	의정부시	호원1동	024A
127	2	경기도	의정부시	송산1동	135A
128	2	경기도	안양시 만안구	안양7동	0681
129	2	경기도	안양시 만안구	안양9동	016A
130	2	경기도	안양시 동안구	달안동	075A
131	2	경기도	안양시 동안구	범계동	015A
132	2	경기도	부천시 원미구	원미1동	0131
133	2	경기도	부천시 원미구	중3동	064A
134	2	경기도	부천시 소사구	소사본2동	0521
135	2	경기도	부천시 소사구	소사본3동	113A
136	2	경기도	광명시	광명2동	0651
137	2	경기도	광명시	광명7동	097A
138	2	경기도	광명시	하안4동	053A
139	2	경기도	평택시	세교동	0101
140	2	경기도	평택시	세교동	062A
141	2	경기도	안산시 상록구	월피동	0501
142	2	경기도	안산시 상록구	성포동	009A
143	2	경기도	안산시 단원구	선부2동	0781
144	2	경기도	안산시 단원구	선부2동	162A
145	2	경기도	고양시 덕양구	화정1동	052A
146	2	경기도	고양시 덕양구	행신3동	133A
147	2	경기도	고양시 일산동구	정발산동	0271
148	2	경기도	고양시 일산서구	일산1동	027A
149	2	경기도	고양시 일산서구	주엽1동	120A
150	2	경기도	구리시	동구동	142A
151	2	경기도	구리시	수택2동	0881
152	3	경기도	남양주시	외부읍	225A
153	3	경기도	남양주시	화도읍	0541
154	3	경기도	남양주시	오남읍	128A
155	2	경기도	남양주시	양정동	017A
156	2	경기도	시흥시	신현동	006A
157	2	경기도	시흥시	정왕 2동	156A
158	2	경기도	시흥시	정왕 1동	0331

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
159	2	경기도	군포시	재궁동	089A
160	2	경기도	의왕시	오전동	096A
161	2	경기도	용인시 처인구	중앙동	0971
162	2	경기도	용인시 기흥구	신갈동	061A
163	2	경기도	용인시 기흥구	어정동	088A
164	2	경기도	용인시 수지구	죽전1동	133A
165	3	경기도	파주시	조리읍	022A
166	2	경기도	파주시	금촌1동	013A
167	3	경기도	이천시	율면	0191
168	3	경기도	김포시	통진읍	030A
169	2	경기도	김포시	김포1동	094A
170	3	경기도	화성시	태안읍	582A
171	2	경기도	화성시	남양동	0721
172	3	경기도	광주시	도척면	0061
173	3	경기도	양주시	남면	027A
174	2	경기도	양주시	양주2동	075A
175	3	경기도	양평군	양동면	0061
176	2	강원도	춘천시	석사동	213A
177	3	강원도	원주시	문막읍	019A
178	2	강원도	원주시	중앙동	0151
179	2	강원도	원주시	단계동	053A
180	2	강원도	강릉시	강남동	071A
181	2	강원도	동해시	북평동	0541
182	2	강원도	속초시	노학동	117A
183	3	강원도	정선군	북평면	0101
184	3	강원도	양구군	양구읍	016A
185	2	충청북도	청주시 상당구	금천동	070A
186	2	충청북도	청주시 흥덕구	사직1동	0241
187	2	충청북도	청주시 흥덕구	산·미·분·장동	147A
188	2	충청북도	청주시 흥덕구	봉명2·송정동	021A
189	2	충청북도	충주시	봉방동	0431
190	2	충청북도	제천시	용두동	038A
191	3	충청북도	옥천군	청산면	0051
192	3	충청북도	음성군	음성읍	027A
193	3	충청남도	천안시	목천읍	063A
194	2	충청남도	천안시	신용동	088A
195	2	충청남도	천안시	쌍용3동	170A
196	3	충청남도	공주시	정안면	0231
197	2	충청남도	공주시	중학동	0411
198	2	충청남도	아산시	온양6동	101A

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
199	3	충청남도	서산시	고북면	004A
200	3	충청남도	금산군	추부면	0471
201	3	충청남도	홍성군	구항면	019A
202	3	충청남도	예산군	대흥면	0011
203	2	전라북도	전주시 완산구	평화2동	055A
204	2	전라북도	전주시 완산구	효자3동	024A
205	2	전라북도	전주시 완산구	효자4동	0661
206	2	전라북도	전주시 덕진구	송천2동	036A
207	2	전라북도	군산시	경암동	0541
208	2	전라북도	군산시	소룡동	070A
209	2	전라북도	익산시	영등1동	093A
210	3	전라북도	남원시	운봉읍	0081
211	2	전라북도	남원시	노암동	013A
212	3	전라북도	무주군	무주읍	013A
213	3	전라북도	부안군	변산면	0151
214	2	전라남도	목포시	용당1동	0531
215	2	전라남도	목포시	하당동	047A
216	2	전라남도	여수시	만덕동	034A
217	2	전라남도	여수시	삼일동	0311
218	2	전라남도	순천시	도사동	002A
219	2	전라남도	광양시	금호동	023A
220	3	전라남도	고흥군	과역면	0041
221	3	전라남도	화순군	화순읍	143A
222	3	전라남도	무안군	청계면	0081
223	3	경상북도	포항시 남구	오천읍	006A
224	2	경상북도	포항시 북구	학산동	0621
225	2	경상북도	포항시 북구	우창동	129A
226	3	경상북도	경주시	안강읍	1781
227	2	경상북도	김천시	대곡동	055A
228	2	경상북도	안동시	용상동	1601
229	3	경상북도	구미시	선산읍	043A
230	2	경상북도	구미시	선주원남동	064A
231	2	경상북도	구미시	양포동	135A
232	2	경상북도	영주시	가흥2동	0011
233	3	경상북도	상주시	공성면	0251
234	2	경상북도	경산시	서부1동	174A
235	3	경상북도	청도군	풍각면	0151
236	3	경상북도	고령군	다산면	027A
237	2	경상남도	창원시	명곡동	1641
238	2	경상남도	창원시	중앙동	107A

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
239	2	경상남도	창원시	가음정동	177A
240	3	경상남도	마산시	내서읍	070A
241	2	경상남도	마산시	오동동	0291
242	2	경상남도	마산시	양덕2동	077A
243	3	경상남도	진주시	문산읍	0041
244	2	경상남도	진주시	상대2동	0431
245	2	경상남도	진주시	가호동	030A
246	2	경상남도	사천시	남양동	0071
247	3	경상남도	김해시	진영읍	031A
248	2	경상남도	김해시	회현동	018A
249	2	경상남도	김해시	활천동	024A
250	3	경상남도	거제시	신현읍	047A
251	3	경상남도	양산시	웅상읍	318A
252	3	경상남도	양산시	동면	0161
253	2	경상남도	양산시	삼성동	005A
254	3	경상남도	함양군	함양읍	0101
255	3	경상남도	함양군	함양읍	091A
256	2	제주도	제주시	이도1동	0281
257	2	제주도	서귀포시	서홍동	0321

< 예비 조사구 목록: 27개조사구>

일련번호	층번호	특광역시	시군구	동 읍 면	조사구
1	1	서울특별시	종로구	혜화동	0211
2	1	서울특별시	중랑구	묵1동	052A
3	1	서울특별시	은평구	역촌2동	0791
4	1	서울특별시	강서구	가양3동	039A
5	1	서울특별시	관악구	신림9동	2761
6	1	서울특별시	강동구	성내1동	088A
7	1	부산광역시	금정구	서4동	0051
8	1	부산광역시	사상구	학장동	126A
9	1	인천광역시	남동구	간석4동	1481
10	1	인천광역시	남동구	논현고잔동	017A
11	1	대전광역시	동구	홍도동	032A
12	2	경기도	수원시 팔달구	우만2동	084A
13	2	경기도	성남시 중원구	상대원2동	0891
14	2	경기도	부천시 원미구	중3동	044A
15	3	경기도	평택시	안중읍	163A
16	2	경기도	고양시 일산동구	정발산동	0821
17	2	경기도	고양시 일산서구	탄현동	091A
18	2	경기도	용인시 수지구	동천동	068A
19	2	충청북도	청주시 흥덕구	강서1동	027A
20	3	충청북도	청원군	내수읍	170A
21	2	충청남도	천안시	신안동	0681
22	3	충청남도	청양군	정산면	0231
23	2	전라북도	정읍시	내장상동	056A
24	3	경상북도	경산시	하양읍	125A
25	2	경상남도	창원시	팔룡동	142A
26	2	경상남도	창원시	사파동	0311
27	3	경상남도	남해군	남해읍	0431