

家口單位 統計調查의
多目的 標本設計

桂 勳 邦

韓國保健社會研究院

머 리 말

家口單位 統計調査는 標本調査區 抽出이 완료되어야만 구체적인 調査計劃을 수립할 수 있으며, 標本調査區는 統計廳에서 5년 주기로 실시하는 人口住宅總調査의 調査區를 이용하여 추출하게 된다.

그런데 조사가 있을 때마다 人口住宅總調査의 調査區를 이용하여 標本調査區를 추출하면 標本과 관련된 모든 業務에 최소한 3個月이 소요되어 제한된 研究期間 내에 調査研究를 수행하는 데에는 標本調査區 抽出 이후의 실제 調査準備 期間이 짧아 어려움을 겪게 된다.

이러한 問題點을 해결하기 위하여 1993년에는 1990年 人口住宅總調査의 調査區 중에서 2,070개의 調査區로 구성되는 400개의 基本標本調査區를 抽出하고 1995년 까지 각종 家口單位 統計調査에 多目的으로 사용하였다.

이번에는 過去의 基本標本調査區 抽出方法을 改善시키고 1995年度 人口住宅總調査의 調査區를 이용하여 1997년부터 2001년까지 5年間 각종 家口單位 統計調査에 多目的으로 사용할 수 있도록 600개 基本標本調査區를 추출하였으며, 여기에는 人口住宅總調査의 調査區 중에서 3,363개 調査區가 포함되었다. 따라서 가구단위 통계조사를 실시할 때마다 이 基本標本調査區 중에서 필요한 數의 實查標本調査區를 抽出하여 사용하면 된다.

本 報告書에서는 基本標本調査區의 抽出節次 및 內容을 자세히 설명하고, 基本標本調査區를 이용하여 實查標本調査區를 抽出하는 方法, 아파트 新築에 따른 追加 標本調査區 抽出方法, 標本推定值 算出方法 등을 事例를 들어 설명하였으므로, 統計學의 기초지식만 있으면 누구라도 이 報告書의 內容에 따라 1997년부터 2001년까지 각종 家口單位 統計調査의 標本設計를 할 수 있을 것으로 期待한다.

특히 앞으로 實施할 1997年度 全國 出産力 및 家族保健實態調査, 1998年度 國民健康 및 保健意識行態調査, 2000年度 障礙人實態調査를 대비하여 事例로 든 標本設計概要는 각 調査의 計劃樹立에도 필요한 것이므로 많이 활용하기 바란다.

研究者는 本 報告書의 原稿를 읽고 유익한 助言을 하여 준 本院의 卞俗榮 博士와 金東奎 博士에게 感謝하고 있으며, 校訂을 맡아 주고 助言도 하여준 高敬煥 主任研究員과 申昌雨 主任電算員에게도 감사하고 있다.

끝으로 本 報告書에 수록한 내용은 어디까지나 著者의 意見일 뿐, 本 研究院의 公式 見解가 아님을 밝혀 둔다.

1996年 12月

韓國保健社會研究院

院長 延 河 清

目 次

I. 序 論	7
II. 標本抽出의 基本 概念	10
1. 人口總調查와 관련된 基本用語	10
2. 標本과 관련된 基本用語	16
3. 確率標本 抽出方法	20
III. 既存 基本標本調查區의 改善點	28
1. 既存 基本標本調查區의 抽出概要	28
2. 改善方向 및 考慮事項	33
IV. 基本標本調查區의 改編	36
1. 基本標本調查區 抽出	36
2. 基本標本調查區의 編制	45
V. 實查標本調查區의 抽出方法	50
1. 基本的인 考慮事項	50
2. 抽出方法	51
VI. 新築 아파트에 대한 措置	61
1. 新築 아파트에 대한 標本調查區 抽出의 必要性	61
2. 新築 아파트 現況資料의 整理	62
3. 新築 아파트의 實查標本調查區 抽出	65
VII. 標本設計 事例	72
1. 1997年度 出産力 및 家族保健實態調查 標本設計概要(案)	72
2. 1998年度 國民健康調查 標本設計概要(案)	83
3. 1998年度 保健意識行態調查 標本設計概要(案)	94
4. 2000年度 障礙人實態調查 標本設計概要(案)	100
VIII. 結論	111

I. 序 論

1. 基本標本調査區 抽出 沿革

1986년 이전까지는 '저소득층 최저생계비 계측조사'와 같은 특수조사를 제외하고는 가구단위 통계조사를 실시할 때마다 통계청의 인구주택총조사(Population and Housing Census)의 조사구를 이용하여 표본조사구를 추출하였다(인구주택총조사는 끝단위가 0 또는 5인 연도에 실시하는 국가 기본통계조사이다).

이와 같이 가구단위 통계조사를 실시할 때마다 매년 인구주택총조사구를 이용하여 표본조사구 추출함으로써 표본조사구의 추출 및 표본조사구 경계선 자료의 복사·정리업무에 3개월 정도가 소요되어 제한된 연구기간 내에 실제 조사연구를 수행하는데에는 조사준비 기간이 짧아 애로가 많았다.

이러한 문제점을 해결하기 위한 하나의 방법으로 4~5년간 각종 가구단위 통계조사에 사용할 수 있도록 많은 수의 표본조사구로 구성되는 기본표본조사구를 추출하여 놓고, 조사가 있을 때마다 이 기본표본조사구 중에서 필요한 수의 표본조사구를 추출하여 사용하는 것이 효율적이다.

그러나 1992년까지는 표본조사구추출을 위한 예산이 각 조사연구사업의 준비단계업무에 포함되어 확보되었기 때문에, 이러한 기본표본조사구의 추출은 예산의 제약으로 불가능한 실정이었다.

다행히 1993년에는 기본표본조사구 추출을 위한 예산이 확보되어 1996년까지 4년간 각종 가구단위 통계조사에 사용할 수 있는 기본표본조사구를 추출한 바 있으며, 이를 각종 가구단위 통계조사의 표본조사구 추출에 사용하였다.

이 기본표본조사구는 1995년 장애인실태조사의 조사구를 마지막으로 모두 사용하였기 때문에 향후 5년간 사용할 수 있는 기본표본조사구의 추출이 시급히 요구되고 있다. 특히 1995년 인구주택총조사의 조사구별 자료를 표본조사구 추출에 사용할 수 있는 단계에 있으므로 이를 이용한 기본표본조사구의 추출이 가능하게 되었다.

2. 研究의 目的

본 연구는 1995년 인구주택총조사의 조사구를 이용하여 향후 5년간 각종 가구 단위 통계조사에 사용할 기본표본조사구를 추출하는데 주된 목적을 두고 있다.

아울러 각종 가구단위 통계조사의 기획단계에서 표본과 관련하여 고려하여야 할 사항과 이 기본표본조사구를 이용한 실사표본조사구의 추출방법을 제시함으로써 표본설계에 경험이 없는 사람이라도 통계학의 기초지식만 있으면 단기간 내에 가구단위 통계조사의 표본설계를 할 수 있도록 하는데 부차적인 목적을 두고 있다.

3. 報告書의 內容

본 보고서는 조사기획 담당자가 표본과 관련하여 고려하여야 할 사항을 이해하여 철저한 조사기획을 수립하는데 도움이 되도록 하고, 표본설계의 기초지식만 있으면 1997년부터 2001년까지 각종 가구단위 통계조사의 표본을 설계할 수 있도록 하는데 주안점을 두고, 내부자료로 작성하였다. 각 장의 내용은 다음과 같다.

제Ⅱ장 : 제Ⅲ장부터 설명되는 내용을 이해하는데 필요한 기본용어와 확률표본 추출방법에 관한 기본개념만을 간단히 설명하였다.

제Ⅲ장 : 1990년 인구주택총조사의 조사구를 이용하여 추출한 기존의 기본표본 조사구의 추출방법은 1995년도 인구총조사의 조사구를 이용하여 추출한 새로운 기본표본조사구의 추출방법을 이해하는데 필수적이므로, 그 추출방법 및 절차를 자세히 설명하였다. 그리고 1995년 인구주택총조사의 조사구를 이용하여 기본표본조사구를 개편할 때에 반영하기 위한 개선점과 고려사항을 정리하였다.

제Ⅳ장 : 1995년 인구총조사의 조사구를 이용하여 개편한 새로운 기본표본 조사구의 추출 내용을 자세히 설명하였다. 따라서 2001년의 기본 표본조사구 개편에서는 여기에 설명한 방법을 그대로 따르거나

개선시키면 된다.

제V장 : 기본표본조사구에서 실사표본조사구를 추출하는 방법을 설명하였으며, 표본추정치 산출공식에는 실사표본조사구가 표본으로 추출된 확률이 고려되므로 실사표본추출과 관련하여 표본추출확률의 산출방법에 대하여도 설명하였다.

제VI장 : 실사표본조사구에 아파트가 신축된 경우에는 신축 아파트는 제외하고 조사를 하여야 하며 그 대신 전국의 신축 아파트를 대상으로 표본조사구를 추출하여 조사에 포함시켜야 하는데, 여기서는 신축 아파트에서 추가되는 표본조사구수 산출방법과 표본조사구 추출절차 및 방법을 예를 들어 설명하였다.

제VII장 : 1997년도 출산력 및 가족보건실태조사, 1998년도 국민건강조사, 1998년도 보건의식행태조사, 2000년도 장애인실태조사의 경우를 예시하여 작성한 표본설계(안)을 사례로 들었다. 국민건강 및 보건의식행태조사의 경우 보건의식행태 부문은 국민건강조사 부문의 조사 대상자 중의 일부를 다시 표본추출하여 조사하며 보고서도 별도로 발간하므로, 별도로 취급하였다.

제VIII장 : 현지 조사과정에서는 실사표본조사구의 관할 동·읍·면이 바뀌어 해당 동·읍·면의 협조를 받을 수 없게 되는 경우, 가구수가 많이 증가한 경우, 조사구 내의 주택이 모두 철거된 경우 등 표본과 관련하여 여러 가지의 문제점이 발생하게 된다. 따라서 이러한 경우를 대비하기 위하여 기본표본조사구의 유지·관리가 필요하게 되므로, 그 방법과 유의사항을 결론으로 제시하였다.

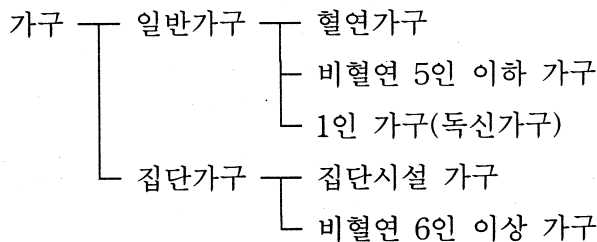
II. 標本抽出의 基本 概念

여기서는 제III장부터 설명되는 내용을 이해하는데 필요한 기본용어와 확률표본추출방법에 관한 기본개념을 간단히 설명하였다.

1. 人口住宅總調査와 관련된 基本用語

가. 家口

家口란 혈연 또는 비혈연을 불문하고 1인 또는 2인 이상이 모여서 취사, 취침 등 생계를 같이하는 생활단위를 말한다. 가구는 크게 一般家口와 集團家口로 나누어지는데, 그 내용은 다음과 같다.



1) 一般家口

가구표본조사에서의 조사대상가구는 일반적으로 일반가구이며, 일반가구의 정의는 조사목적에 따라 조금씩 다르게 되지만, 인구총조사에서 사용한 일반가구의 정의를 기본으로 하고 있다.

가) 血緣家口

통상 가족단위로 이루어져 생활을 같이 하고 있는 가구를 말하며, 다음과

같은 2가지 경우가 포함된다.

- 혈연가족 관계로만 구성된 가구
- 혈연가족과 5인 이하의 비혈연자가 함께 살고 있는 가구(비혈연자가 6인 이상인 경우에는 비혈연자만을 분리하여 집단가구로 취급)

나) 非血緣 5人 이하 家口

비혈연 5인 이하의 사람들만이 함께 생활하는 가구(6인 이상의 경우 집단가구로 분류)

다) 1人 家口(獨身家口)

혼자서 생활하는 가구

2) 集團家口

가구표본조사에서는 집단가구는 특수한 경우를 제외하고는 조사대상가구에서 제외되지만, 일반가구와 구별을 명백히 하기 위하여 그 개념을 알아 둘 필요가 있다.

가) 集團施設 家口

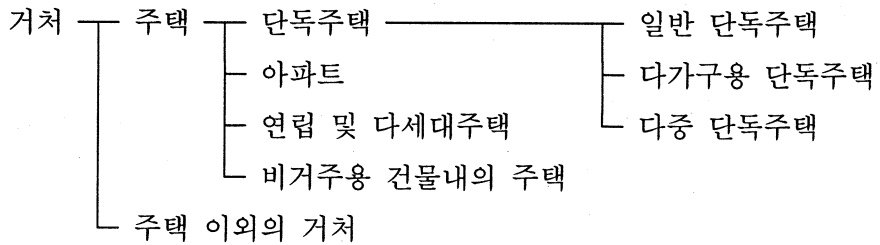
- 기숙사, 수녀원, 사찰 등에서 거주하는 사람들의 경우
- 고아원, 양로원, 요양소 등의 특수사회시설에 2인 이상이 수용되어 있는 사람들의 경우(인원수가 1명인 경우는 일반가구의 1인가구로 취급하지만 2인 이상이면 그 명칭을 위주로 집단시설가구로 한다)

나) 非血緣 6人 이상 家口

혈연관계가 없는 6인 이상의 사람들이 동일한 거처 내에서 생활을 같이 하고 있는 경우(20인 이상인 경우 집단시설 가구로 분류)

나. 居處

居處(living quarter)란 사람이 살고 있는 모든 장소를 통칭하는 말로서, 구조적으로 분리되고 독립된 하나의 거주단위를 말하며, 다음과 같이 구분한다.



1) 住宅

住宅이란 사람이 살 수 있도록 거주목적으로 지어진 집으로서, 다음의 요건을 갖춘 건물을 말한다.

- 영구적 성격의 건물
- 한개 이상의 방과 부엌
- 독립된 출입구
- 관습상 소유 또는 매매의 1단위

가) 單獨住宅

단독주택에는 다음과 같은 주택이 포함된다.

- 一般 單獨住宅 : 원칙적으로 한 가구가 생활할 수 있도록 건축된 주택을 말하며, 주택 내의 일부를 다른 가구에 세를 주는 경우가 있다.
- 다가구용 단독주택 : 여러 가구가 살 수 있도록 지어진 하나의 건물에 가구별로 독립적인 생활이 가능하도록 화장실, 부엌, 출입구가 각각 독립되어 있으나 각 단위별로 소유 또는 매매가 불가능한 주택을 말한다. 다가구용 단독주택은 건물 전체를 하나의 주택으로 본다.
- 多重 單獨住宅 : 여러 가구가 살 수 있도록 지어진 하나의 주택에 가구별로 독립적인 생활이 가능하도록 부엌, 출입구가 각각 독립되어 있지만 화장실을 공동으로 사용하며, 가구 단위별로 소유 또는 매매가 불가능한 단독주택을 말한다. 다중 단독주택의 경우에도 건물 전체를 하나의 주택으로 본다.

나) 아파트

한 건물내에 여러 가구가 거주할 수 있도록 지어진 5층 이상의 건물로서,

구조적으로 한 가구씩 독립하여 살 수 있도록 구분된 주택을 말한다. 아파트의 경우 구분된 각 주거단위가 하나의 주택이 된다(1개 동 전체가 하나의 주택이 아님).

다) 聯立 및 多世帶住宅

한 건물 내에 여러 가구가 살 수 있도록 건축된 4층 이하의 건물로서 건축당시 연립주택 또는 다세대주택으로 허가받은 주택을 말한다.

- 聯立住宅 : 연립주택의 경우에도 각 주거단위가 하나의 주택이 된다.
- 多世帶住宅 : 건물 연면적이 166m² (200평)이하이며, 연립주택의 경우와 같이 구분된 각 단위별로 분리하여 등기가 된다. 다세대주택의 경우 연립주택과 같이 구분된 각 주거단위가 하나의 주택이 된다.

라) 非居住用 建物 内の 住宅

거주 목적이 아닌 건물 내에 사람이 거주할 수 있도록 구획된 부분으로서, 주택의 요건을 갖추고 있는 경우를 말한다(예: 상가 건물 내의 주택).

2) 住宅 이외의 居處

주택의 요건을 갖추지 못하고 있는 건물로서 사람이 살고 있는 장소를 말하며, 다음과 같은 것이 포함된다.

- 오피스텔(사무실 겸용주택)
- 호텔·여관 등 숙박업소의 객실
- 기숙사 및 특수사회시설(기숙사, 고아원, 수녀원, 양로원, 보육원 및 대사찰 등과 같이 집단수용을 위한 구조물)
- 기타 임시적 거주를 위한 구조물(천막집, 판잣집, 비닐하우스, 움막, 암자, 건설공사장의 임시막사 등)

다. 調查區

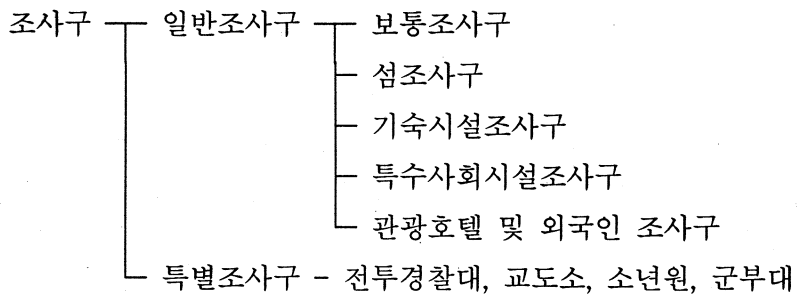
인구주택총조사의 실시를 위하여 전국의 행정 동·읍·면을 각각 일정한 기준에 의하여 여러 개의 조사구역으로 분할하였으며, 각 조사구역을 조사구라 한다.

조사구는 일차적으로 조사원의 조사담당 구역을 명확히 하여 가구나 주택의 누락 또는 중복을 방지하고 조사원의 업무량을 적정하게 배분하기 위하여 설정한다.

그리고 인구총조사가 완료된 후에는 각종 가구표본조사의 표본조사구 추출의 기본자료를 제공하고 소지역별 통계작성에 기초자료로 활용하기 위하여 설정하는 것이다.

1) 調査區의 種類

조사구는 일반조사구와 특별조사구로 분류되며, 그 내용은 다음과 같다.



2) 一般調査區

일반조사구란 조사원이 가구를 방문하여 조사활동을 함에 있어 특별한 출입 제한이 없이 응답자와의 면접이 자유로운 지역에 일정 기준에 따라 설정한 조사구이다.

가) 普通調査區(특성번호 1)

- 일반가구와 비혈연자 6인 이상 19인 이하가 거주하는 집단가구를 대상으로 주요 지형지물 또는 통·리 경계를 이용하여 설정한 조사구
- 1개 조사구당 평균 60가구를 기준으로 설정함을 원칙으로 하되 지리적 조건 등으로 인하여 기준 가구수 적용이 어려운 경우에 한하여 50~70 가구 범위내에서 설정
- 독신자 아파트(사원아파트, 공무원 아파트 등), 오피스텔, 가족이 거주하는 사원주택 등은 보통조사구로 설정

나) 섬 調査區(특성번호 2)

- 육지와 완전히 분리되어 4면이 바다로 둘러 싸인 섬을 대상으로 보통 조사구와 동일한 기준에 의하여 설정하는 조사구
- 유인도 또는 무인도가 군집되어 있는 섬지역은 거주 가구수와 거처수를 고려하여 여러 개의 섬을 묶어 1개의 섬조사구로 설정하되, 해상교통이 불편하여 업무에 막대한 지장을 초래할 우려가 있는 섬에 한하여 기준가구수(60가구)에 미달되더라도 1개 섬조사구로 설정

다) 寄宿施設 調査區(특성번호 3)

- 기숙시설의 명칭 유무를 불문하고 회사, 공장, 학교, 병원, 수녀원, 대사찰 및 기타 단체 등의 소속 직원, 학생 등 비혈연자 20인 이상이 공동으로 숙식을 할 수 있도록 만들어진 시설단위별로 설정하는 조사구
- 비혈연자 20인 이상이 공동으로 숙식을 할 수 있도록 만들어진 시설로서 조사일 현재 20인 미만이 거주하고 있는 경우도 기숙시설 조사구로 설정
- 단독주택 등에서 하숙업을 하거나 기숙사와 동일한 형태로 실제 사용되고 있는 경우, 주인가구를 제외하고 비혈연자 20인 이상이 거주하고 있으면 기숙시설 조사구로 설정(단, 주인가구는 보통조사구에 포함)
- 기숙시설 내에 거주 인원이 300인을 초과할 경우에는 300인을 기준으로 동(棟), 층 또는 호·실 단위로 기숙시설을 분할하여 조사구를 설정

라) 特殊社會施設 調査區(특성번호 4)

- 특수사회시설 조사구란 고아원, 양로원, 육아원, 요양원, 나환자 수용소 등과 같이 일정 장소에 비혈연자 2인 이상이 공동으로 숙식을 할 수 있도록 설치된 시설단위별로 설정하는 조사구
- 특수사회시설 조사구로 설정되는 수용시설은 다음과 같음.
 - ㉠ 아동복지법에 의하여 설치된 수용시설
육아원, 보육원, 희망원, 영아원, 교호시설, 자립지원시설 등
 - ㉡ 노인복지법에 의하여 설치된 수용시설
양로원, 노인요양원, 노인복지주택 등

㉔ 장애인복지법에 의하여 설치된 수용시설

맹·농아시설, 정신장애자시설, 지체부자유자 등의 요양시설

㉕ 특수병원, 수용소

정신병원, 결핵요양원, 나환자수용소, 전염병 환자만을 수용하는 시설 등

㉖ 기타

윤락여성 직업보도를 위한 수용시설, 그 밖에 사실 기도원 등 ㉑~㉕에 준하는 수용시설

- 나환자의 집단 부락은 보통조사구(특성번호 1)로 설정

마) 觀光호텔 및 外國人 調査區(특성번호 5)

- 觀光호텔 및 外國인 調査區란 外國인이 투숙한 觀光호텔과 20인 이상의 外國인이 집단으로 거주하는 건물 또는 10가구 이상의 外國인 가구가 모여서 생활하고 있는 지역을 대상으로 설정하는 조사구
- 觀光호텔은 3등급(무궁화 2개) 이상에 한하여 호텔단위별로 觀光호텔 조사구로 설정하고 3등급 미만(무궁화 1개)의 觀光호텔, 여관 등은 보통조사구에 포함하여 설정
- 20인 미만의 外國인 또는 10가구 미만의 外國인 가구가 집단으로 거주하거나 韓國인 가가와 섞여서 거주하는 지역은 보통조사구에 포함하여 설정

2. 標本과 관련된 基本用語

가. 統計調査와 母集團

統計調査(statistical survey)란 통계작성을 목적으로 자료를 수집하는 조사를 말한다. 어떤 통계조사든지 자료를 얻고자 하는 대상이 되는 단위가 정해져야 하는데, 이 단위를 基本單位(elementary unit)라 한다. 기본단위를 때에 따라서는 元素(element)라고도 한다.

일반적으로 기본단위는 사람, 가구, 농가, 사업체 등이지만, 이와는 달리 전산 화일의 단위 레코드일 수도 있다. 특히 하나의 통계조사에서 2개 이상의 기본단위가 있을 수 있다. 예를 들면, 가구와 사람, 농가와 경지면적 등의 경우이다.

통계학에서는 기본단위의 전체 집단을 母集團(population)이라 하는데, 모집단이란 용어는 학문적인 용어이므로 일반인이 그 뜻을 알지 못하는 경우가 있기 때문에 통계조사 실무에서는 일반인이 쉽게 이해할 수 있는 調查對象集團 또는 간단히 調查對象이란 용어를 사용하고 있다.

나. 全數調查와 標本調查

통계조사는 조사대상이 되는 모든 단위에 대하여 자료를 수집하는 경우와 조사대상이 되는 단위 중의 일부를 추출하여 자료를 수집하는 경우가 있는데, 조사대상이 되는 모든 단위에 대하여 자료를 수집하는 경우를 全數調查(complete enumeration)라 하며, 전체 단위 중에서 일부의 단위를 추출하여 자료를 수집하는 경우를 標本調查(sample survey)라 한다. 표본조사를 위하여 모집단의 단위 중에서 추출된 일부의 단위를 標本(sample)이라 한다.

통계조사를 전수조사로 실시하든 표본조사로 실시하든 모집단은 조사목적에 따라 ① 시점(또는 대상 기간), ② 지역, ③ 단위, ④ 범위 등이 명백하게 정의되어야 한다. 인구총조사의 경우를 예로 들면 다음과 같다.

- ① 시점 : 1995년 11월 1일 0시 현재
- ② 지역 : 대한민국에
- ③ 단위 : 거주하는 사람으로 외국인도 포함되나
- ④ 범위 : (1) 외국 외교관 및 수행원, (2) 공무로 국내에 체류중인 외국인, (3) 국제연합 소속기관의 외국인 직원 및 수행원, (4) 국내 주둔 외국 군인 및 군속과 (5) (1)~(4)까지의 가족은 제외한다. 이와 반대로 외국 주재 우리나라 공관원 및 그 가족은 포함하되, 해외 취업 및 유학 중인 한국인은 국내 거주자가 아니므로 제외한다.

이런 면에서 모집단은 잘 定義된 集合(well-defined set)이어야 하며, 표본은 모집단의 部分集合(subset of population)이 되는 것이다.

다. 母數와 統計量

모집단에서 직접 산출된 총량, 평균, 구성비, 표준편차, 분산 등의 각종 통계치를 통틀어 母數(parameter)라 한다. 통계학 분야에서는 모수라는 용어를 주로 사용하지만 통계조사 분야에서는 모수라는 용어 대신에 母集團의 統計値라는 용어를 주로 사용한다.

이에 대하여 표본에서 산출된 각종 통계치를 統計量(statistic)이라 한다. 통계학 분야에서는 통계량이란 용어를 주로 사용하지만, 통계조사 분야에서는 標本推定値(sample estimate)라는 용어를 주로 사용한다.

표본조사의 궁극적인 목표는 모집단의 통계치에 대한 표본추정치를 얻는데 있으며, 이것은 표본에서 얻은 추정치를 사용하여 모집단의 특성을 설명하게 되기 때문이다.

라. 標本設計

모집단에서 표본을 추출하는 것을 標本抽出(sample selection)이라 하며, 표본추정치를 산출하는 것을 推定(estimation)이라 한다. 그리고 표본추출과정과 추정과정을 통합하여 標本設計(sample design)라 한다.

일반적으로 표본추출과정을 이해하는데 어려움이 별로 없지만 표본추정치를 산출하는 과정은 확률의 개념이 토대가 되므로 통계학의 기초지식이 요구된다.

마. 確率標本과 非確率標本

표본을 추출할 때에는 기본적으로 모집단의 모든 기본단위가 표본으로 추출되는 기회를 갖도록 하여야 한다. 즉 표본으로 추출되는 확률이 '0'이 되는 기본단위가 있어서는 아니된다. 그리고 각 기본단위가 표본으로 추출되는 확률은 같을 수도 있고 다를 수도 있지만, 그 확률을 알 수 있어야 한다.

이와 같이 표본을 추출하는 과정에서 확률에 근거하여 표본을 추출하면, 이 표본을 確率標本(probability sample)이라 한다.

표본을 추출하는 과정에서 확률을 무시하고 표본을 추출하면, 이 표본을 非確率標本(non-probability sample)이라 한다.

확률표본을 사용하는 이유는 ① 표본을 추출하는 사람의 편견이나 주관을 배제할 수 있고, ② 모수와 표본오차를 추정할 수 있는 공식이 존재하기 때문이다.

바. 標本誤差와 非標本誤差

전수조사의 통계치와 표본에서 얻은 추정치는 일반적으로 어느 정도의 차이가 나게 마련이며, 이 차이를 標本誤差(sampling error)라 한다. 표본오차는 표본을 추출하기 때문에 그 과정에서 어쩔 수 없이 우연적으로 발생하는 오차이며, 표본을 추출하는 것과는 관계없이 자료수집과정이나 자료처리과정에서 발생하는 오차, 예를 들면 응답자가 거짓으로 응답하는 경우, 조사원이 조사표에 조사내용을 잘못 기입하는 경우, 자료입력이 잘못되는 경우 등에 의한 오차를 非標本誤差(non-sampling error)라 한다.

따라서 표본오차는 표본조사에서만 발생하지만, 비표본오차는 전수조사와 표본조사에서 모두 발생한다. 즉 전수조사에서는 비표본오차만 발생하지만 표본조사에서는 비표본오차와 표본오차가 함께 발생한다.

전수조사에서는 비표본오차만 발생하고 표본조사에서는 비표본오차와 표본오차가 함께 발생하기 때문에 전수조사 결과가 표본조사 결과보다 정확하다고 생각할 수도 있지만, 표본조사 결과가 더 정확한 경우도 있다. 이에 관한 대표적인 예로서 1950년 미국의 인구총조사(Population Census)에서 수집된 노동력에 관한 통계자료와 2만 가구의 표본에 대하여 실시한 노동력 표본조사 결과를 들 수 있다.

1950년 인구총조사 결과가 발표되었을 때, 취업자수 및 실업자수가 노동력 표본조사에서 추정된 통계치와 너무나 차이가 난다는 것이 판명되었다. 처음에는 이 차이가 표본오차에 기인한다고 생각되었으나 사실은 그렇지 않았다. 즉 인구총조사에서의 응답문제는 매월 조사의 표본오차보다 더 큰 오차를 내었고, 이것은 대부분이 조사에 대한 경험이 없는 사람을 조사원으로 채용한데 그 원인이 있었다. 그러므로 인구총조사 자료의 이용자들에게 믿을 수 있는 노동력통계로는 표본조사의 결과를 이용할 것을 권고하였다.

표본오차와 관련된 용어로 精度(precision)란 것이 있는데 표본의 정도가 높다

는 것은 표본오차가 작다는 뜻이며, 표본의 정도가 낮다는 것은 표본오차가 크다는 뜻이다.

3. 確率標本 抽出方法

가. 單純任意抽出法

單純任意抽出法(simple random sampling)은 가장 기본이 되는 표본추출방법으로, 동일한 조건에서 제비뽑는 방법으로 표본을 추출하는 방법이다. 여기서 동일한 조건이란 모집단익 각 기본단위가 표본으로 추출되는 기회가 서로 같다는 뜻이다. 또한 전체 제비의 수와 표본으로 추출된 제비의 수를 알 수 있으므로, 표본으로 추출된 제비가 어떤 확률로 추출되었는지는 알 수가 있다.

단순임의표본추출법에서 실제적으로 제비를 만들어 표본을 추출하는 경우는 거의 없으며, 기본적으로 亂數表(random number table)를 사용한다.

난수표는 0~9까지의 숫자가 어떤 일정한 순서나 아무런 의미도 없이 무질서하게 나열되어 있는 수표이다. <表 II-1>은 난수표의 예를 든 것이다. 난수표는 일반적으로 보기 쉽게 하기 위하여 5단위씩 묶어서 정리하고 있다.

<表 II-1> 亂數表

줄번호	(1)	(2)	(3)	(4)	(5)	(6)	(7)
①	23418	23508	91057	76455	54941	72711	39406
②	57304	73678	08273	62941	02349	71389	45605
③	77643	98489	86448	73652	98210	43546	27174
④	68326	65613	01635	07607	11826	91326	29624
⑤	64472	72294	95596	81009	96810	17100	35065
⑥	88205	37913	89273	10329	81060	33449	68056
⑦	98455	78685	71250	59243	56135	80647	51304
⑧	75807	45794	96317	27615	18432	19473	64258

이 난수표를 이용하여 60명의 학급에서 12명을 표본으로 추출하는 경우를 예로 들면 다음과 같다.

- (1) <表 II-1>의 난수표에서 아무 곳에서나 출발점을 정한다. 여기서는 2번째 줄의 3번째 칸에 있는 '08273' 중에서 가운데 수 '2'를 출발점으로 택했다고 하자(출발점은 아무렇게나 택하면 되는데, 여기서는 눈을 감고 연필로 찍었는데 이 수가 선택되었다고 가정한다. 따라서 표본을 추출하는 사람에 따라 다른 수가 택해 진다).
- (2) 60명은 2단위의 수이므로 출발점부터 2단위의 수를 아래로 읽어 내려 간다. 그러면 출발점의 2단위 수는 '27'이므로 27번의 학생을 표본으로 추출한다(여기서는 출발점부터 아래로 내려갔지만 위쪽으로, 또는 좌우의 어느 쪽으로 가더라도 2단위의 수를 택하기만 하면 된다).
- (3) 그 아래의 2단위 수는 '44'이므로 44번 학생을 표본으로 추출한다.
- (4) '44' 아래의 2단위 수는 '63'인데, 학생은 60번까지만 있으므로 '63'은 무시하고, 그 다음의 2단위 수가 '59'이므로 59번 학생을 표본으로 추출한다.
- (5) '59' 다음의 수는 '27'인데 27번 학생은 첫 번째로 이미 추출되었으므로 제외하고, 다음의 2단위 수가 '25'이므로 25번 학생을 표본으로 추출한다.
- (6) 이와 같은 과정을 반복하여 12명을 추출하면 된다.

이와 같이 표본을 추출하면 표본을 추출하는 사람의 주관이 작용하지 않기 때문에, 단순임의추출법을 單純無作爲抽出法이라고도 한다.

모집단의 단위의 수가 N 개이고 표본의 단위수가 n 개인 경우에 표본의 크기(n)을 모집단의 크기(N)로 나눈 ' n/N '을 標本抽出率(sampling ratio, sampling fraction)이라 한다. 앞의 예에서 $N = 60$ 이고 $n = 12$ 이므로 표본추출률은 $12/60 = 1/5$ 이다.

표본추출 목적의 난수표는 쉽게 말해서 0~9까지의 10개의 숫자가 각각 2번씩 표시된 20면체의 주사위를 계속하여 굴려서 맨 윗면에 나오는 숫자를 기입하여 놓은 표라고 생각하면 된다.

이 주사위가 정확한 규격으로 만들어 진 것이라면, 0~9까지의 각 숫자가 나올 확률은 $1/10$ 로 같게 된다. 확률이 같다는 것은 10번을 굴리는 경우에 0~9까지의 각 숫자가 꼭 한번씩 나온다는 뜻이 아니라, 이 주사위를 많이 굴리면 굴

릴수록 각 숫자가 나오는 횟수는 점점 더 일치하게 된다는 뜻이다. 다른 말로 표현하면 주사위를 굴리는 횟수가 증가할수록 각 숫자의 출현률은 $1/10$ 에 점점 더 접근한다는 뜻이다. 근래에는 컴퓨터에서 소프트웨어를 이용하여 이러한 난수를 만들어 낼 수가 있다. 휴대용 계산기 중에도 난수를 만들어 내는 기능이 있는 것도 있다.

나. 系統抽出法

系統抽出法(systematic sampling)은 60명의 학급에서 12명을 표본으로 추출할 때에, 출석부에서 2번, 7번, 12번 등의 학생을 표본으로 추출하는 방법이다. 이 경우에 60명 중에서 12명을 추출하므로 5명마다 1명을 추출하게 된다.

여기서, $60 \div 12 = 5$ 를 抽出間隔(sampling interval)이라 한다. 그리고 맨 처음의 번호를 出發番號(starting number)라 하며, 출발번호는 추출간격 이하의 수를 난수표에서 택한다. 따라서 추출간격 5 이하의 수로 2가 추출되었다면, 2번부터 매 5번째 번호의 학생을 표본으로 추출하면 된다. 이와 같이 추출된 표본을 系統標本(systematic sample)이라 한다.

그런데 60명중에서 8명의 계통표본을 추출하려면 추출간격은 $60 \div 8 = 7.5$ 가 되어 곤란하게 된다.

이 때에는, 추출간격 7.5 이하의 수를 난수표에서 택한다. 그러나 난수표에는 소수점을 가진 수는 없으므로 75 이하의 수를 택한다. 여기서 75 이하의 36이 선택되었다면 출발번호는 3.6이 된다.

다음에는 이 출발번호 3.6에 추출간격 7.5를 60을 초과할 때까지 계속 누적한다. 그러면 3.6 , $3.6 + 7.5 = 11.1$, $11.1 + 7.5 = 18.6$, \dots , $41.1 + 7.5 = 48.6$, $48.6 + 7.5 = 56.1$, $56.1 + 7.5 = 63.6$ 이 된다. 계산결과 9개의 수 3.6, 11.1, 18.6, 26.1, 33.6, 41.1, 48.6, 56.1, 63.6에서 각각 소수점 이하의 수를 무시하고 3번, 11번, 18번, 26번, 33번, 41번, 48번, 56번의 학생을 표본으로 추출하면 된다 (60번 학생까지만 있으므로 63번은 제외한다). 여기서, 3, 11, \dots , 48, 56을 표본 지정번호(sample designation number)라 한다.

모집단의 기본단위가 거의 임의의 순서로 나열되어 있다면, 계통추출의 결과

는 단순임의추출의 결과와 상당히 일치하게 되며, 계통추출법은 단순임의추출법보다 표본추출절차가 간단하기 때문에 실제로는 단순임의추출법보다 계통추출법을 많이 사용한다.

다. 層化抽出法

60명의 학생중에서 20명이 여학생이며, 출석부의 번호는 남녀의 구별없이 신장의 크기순으로 부여되었다고 하자. 이러한 상태에서 12명을 표본으로 추출할 때, 단순임의추출법이나 계통추출법을 사용하면 표본으로 추출된 12명 모두가 남학생일 수도 있고, 이와는 반대로 12명 모두가 여학생일 수도 있게 된다. 표본은 모집단의 각종 특성이 가급적 잘 반영되는 것이 가장 바람직한데, 이와 같이 모집단의 특성이 표본에 제대로 반영되지 않고 차이가 나게 된다면 이러한 표본은 잘 추출된 표본이라고 할 수 없다.

앞에서와 같은 경우에 우선 남학생과 여학생의 2개 집단으로 나누어 표본으로 추출하려는 12명을 각 집단의 학생수에 비례하도록 배분하면, 남학생 집단에는 $12 \times (40/60) = 8$ 명, 여학생 집단에는 4명이 배분된다. 따라서 40명의 남학생 집단에서 8명을 표본으로 추출하고, 20명의 여학생 집단에서 4명을 표본으로 추출하면 된다. 이 때, 각 집단에서는 단순임의추출법이나 계통추출법으로 표본학생을 추출한다.

이와 같은 표본추출방법을 層化抽出法(stratified sampling)이라 한다. 층화추출법에서 모집단을 여러 개의 집단으로 분류하는 것을 層化(stratification)한다고 하며, 이들 집단을 層(strata)이라 한다. 층화를 할 때에는 층 내에서는 동질적이 되도록 하고, 층과 층 사이에는 이질적이 되도록 하여야 한다.

층화의 지표로는 조사하고자 하는 내용과 상관관계가 큰 변수를 지표로 사용하여야 한다. 예를 들면, 체중을 조사하고자 하는 경우에 체중과 신장이 상관관계가 있고 체중과 앉은 키도 상관관계가 있을 때, 체중과 신장의 상관관계가 체중과 앉은 키의 상관관계보다 크다면 상관관계가 큰 신장으로 층화하여야 한다. 이 때 신장으로 먼저 층화하고, 각 층에서 다시 앉은 키로 층화할 수도 있다.

각 층에서 단순임의추출법으로 표본을 추출하면 層化單純任意抽出法이 되며,

계통추출법으로 표본을 추출하면 層化系統抽出法이 된다.

층화추출법을 사용하는 이유는 기본적으로 표본오차를 감소시키기 위하여 사용하며, 일반적으로 층화를 하지 않는 경우보다 표본오차가 감소하게 된다.

라. 集落抽出法

2개 이상의 기본단위로 구성되는 집합체를 集落(cluster)이라 하는데, 집락을 표본으로 추출하는 경우를 集落抽出法(cluster sampling)이라 한다.

인구조사의 경우에 가구를 표본으로 추출하면 가구는 가구원의 집합체이므로 집락이 되며, 가계조사의 경우에 가구를 표본으로 추출하면 가구는 기본단위가 된다. 또한 근로자의 건강의식행태조사의 경우에 사업체를 표본으로 추출하면 사업체는 근로자의 집합체이므로 집락이 되며, 생산량조사의 경우에 사업체를 표본으로 추출하면 사업체는 기본단위가 된다.

표본집락은 단순임의추출법이나 계통추출법으로 추출할 수가 있으며, 층화단순임의추출법이나 층화계통추출법으로도 추출할 수가 있다.

집락표본추출법은 기본단위의 명부가 없을 때와 기본단위의 명부가 있더라도 기본단위를 표본으로 추출하면 표본이 각 지역에 골고루 분포되어 예산과 조사기간이 많이 소요되는 경우에 사용한다.

집락의 개념을 지역단위로 확대하여 집락 대신에 통·리, 반, 인구총조사의 조사구 등과 같은 지역단위를 표본으로 추출하는 경우가 많은데, 이러한 방법을 地域抽出法(area sampling)이라 한다. 일반적인 대규모 가구단위 통계조사에서는 지역표본을 사용하고 있다.

마. 多段抽出法

표본집락이나 표본지역 내에서 다시 더 작은 단위를 표본으로 추출하는 경우를 2段抽出法(two-stage sampling)이라 한다.

2단추출법의 경우에 처음 단계에서 표본으로 추출하는 단위를 1次抽出單位(primary sampling unit)라 하며, 표본으로 추출된 1차추출단위 내에서 다시 표본으로 추출되는 단위를 2次抽出單位(secondary sampling unit)라 한다.

국민학교를 표본으로 추출하고 표본으로 추출된 국민학교에서 다시 학생을 표본으로 추출한다면, 1차추출단위는 국민학교이고 2차추출단위는 학생이 된다.

통·리를 표본으로 추출하고, 표본으로 추출된 통·리에서 다시 반을 추출한 다음, 표본으로 추출된 반에서 다시 가구를 표본으로 추출한다면, 이 추출방법은 3段抽出法(three-stage sampling)이 된다. 이 때, 1차추출단위는 통·리, 2차추출단위는 반, 3차추출단위는 가구가 된다.

이와 같이 여러 단계로 표본을 추출하는 방법을 통틀어 多段抽出法(multi-stage sampling)이라 한다. 따라서, 다단추출법에서 1차추출단위는 반드시 집락이거나 지역이어야 한다.

바. 確率比例抽出法

집락이나 지역단위를 표본으로 추출할 때에 집락이나 지역단위를 구성하는 단위의 수에 비례하는 확률로 표본집락을 임의추출하는 방법이 있다.

어떤 통의 10개 반에서 3개 반을 추출한다고 할 때, 이 방법으로 표본반을 추출하는 경우의 예로 <表 II-2>의 '방법 1'을 들기로 한다.

<表 II-2> 確率比例 標本抽出 作業表

기초자료		가구 일련번호	방법 1		방법 2	
반	가구수		표본 지정번호	표본 추출확률	표본 지정번호	표본 추출확률
1	25	1~25				
2	20	26~45			26	20/88
3	35	46~80	48	35/264		
4	17	81~97				
5	39	98~136	108	39/264	114	39/88
6	28	137~164				
7	16	165~180				
8	24	181~204			202	24/88
9	38	205~242				
10	22	243~264	250	22/264		

〈表 II-2〉에서 1반부터 차례로 각 가구에 일련번호를 부여하면, 1반은 25가구로 1~25번까지 부여됨으로 1반의 일련번호란에는 1~25를 기입하고, 2반은 20가구로 26번부터 45번까지 부여됨으로 26~45를 일련번호란에 기입한다.

이와 같은 방법으로 10반까지 일련번호를 기입하면 각 반의 마지막 가구번호는 누적가구수와 같게 되며, 최종 누적가구수 264는 총가구수와 일치하게 된다.

다음에는 최종 가구번호 264이하의 3개 수를 난수표에서 찾는다. 48, 250, 108의 3개 수가 선정되었다면 48번, 250번, 108번이 표본지정번호가 된다. 첫 번째 지정번호 48번이 속한 3반의 표본지정번호에 48을 기입한다. 그리고 마찬가지로 방법으로 250번 가구가 속한 10반과 108번 가구가 속한 5반의 표본지정번호에 각각 250과 108을 기입한다. 그러면 표본지정번호가 기입된 3반, 5반, 10반이 표본반이 된다.

이 때, 각 반이 표본으로 추출되는 확률은 [각 반의 가구수 / 총 가구수]가 된다. 따라서 표본으로 추출된 3반의 가구수는 35가구이므로, 이 3반이 표본으로 추출된 확률은 $[35/264]$ 가 되며, 5반과 10반의 경우에는 각각 $[39/264]$ 및 $[22/264]$ 가 된다. 여기서, 각 반의 표본추출확률은 가구수에 비례한다는 것을 알 수 있다.

이와 같은 추출방법을 確率比例抽出法(sampling with probability proportional to size, PPS sampling)이라 한다.

이와는 달리 가구수에 비례하는 확률로 계통추출하는 방법이 있으며, 이 방법에 대하여는 〈表 II-2〉의 ‘방법 2’를 예로 들기로 한다. 이 방법에서는 최종 누적가구수 264를 추출하고자 하는 표본반의 수 3으로 나누어, 추출간격 88을 구한다.

그리고 88 이하의 난수 하나를 택하여 추출번호로 한다. 이 수가 26이라면, 26번 가구는 2반에 속하므로 2반의 표본지정번호에 26을 기입한다. 다음에는 $26(\text{출발번호}) + 88(\text{추출간격}) = 114$ 이므로 114번 가구가 속하는 5반의 표본지정번호에 114를 기입한다. 같은 방법으로 $114 + 88 = 202$ 이므로 202번 가구가 속한 8반의 표본지정번호에 202를 기입한다.

그러면 표본지정번호가 기입된 2반, 5반, 8반이 표본반이 된다. 이 때에, 각 반이 표본으로 추출되는 확률은 [가구수 / 추출간격]이므로, 2반이 표본으로 추

출된 확률은 $[20/88]$ 이 되며, 5반과 8반의 경우에는 각각 $[39/88]$ 및 $[24/88]$ 가 된다. 여기서도, 각 반의 표본추출확률은 가구수에 비례한다는 것을 알 수 있다.

이와 같이 확률비례추출법에 계통추출법이 추가된 방법을 確率比例系統抽出法 (systematic sampling with proportional to size, PPS systematic sampling) 이라 한다.

Ⅲ. 既存 基本標本調査區의 改善點

여기서는 1990년 인구총조사의 조사구를 이용하여 1993년도에 추출한 기존 기본표본조사구의 추출방법과 절차를 자세히 설명하였다.

그리고 1995년 인구주택총조사의 조사구를 이용하여 기본표본조사구를 개편할 때에 반영하기 위하여 실사표본조사구를 추출하는 과정에서의 개선점과 기타 고려사항을 정리하였다.

1. 既存 基本標本調査區의 抽出概要

가. 企劃段階에서 考慮된 事項

1) 基本標本調査區 抽出 對象地域

기본표본조사구는 전국을 대상지역으로 하여 추출하는 것이 바람직하지만, 교통이 불편한 섬지역에서 표본조사구가 추출되면 조사비용이 많이 소요될 뿐만 아니라 실제로 현지조사가 곤란하게 되는 경우도 있으므로, 각 시·도의 섬지역은 기본표본조사구 추출 대상지역에서 제외시키기로 하였다.

실제적으로 각 시·도의 섬지역의 총가구수는 전국 가구수의 1% 미만이므로, 이들 가구를 제외시켜도 조사결과에 별다른 영향을 미치지 않는다.

2) 基本資料

가구단위 통계조사의 표본조사구 추출을 위한 기본자료로는 통계청에서 매 5년마다 실시하고 있는 인구주택총조사의 조사구 자료뿐이다. 따라서 1993년 현재에는 1990년 인구주택총조사의 조사구별 자료를 이용하여 표본조사구를 추출

하게 된다.

이 조사구의 조사구당 평균 가구수는 65가구 정도이므로 특별한 경우를 제외하고는 조사관리의 면에서 적절한 규모라 볼 수 있다.

그리고 장애인실태조사의 경우와 같이 1개 조사구에서 200가구 정도를 조사하여야 하는 경우에는 인구주택총조사의 조사구를 3~4개씩 통합하여 하나의 조사구로 취급하면 된다.

그리고 가구단위 통계조사의 표본조사구를 추출 하는 것이므로, 인구주택총조사의 조사구 중에서 기숙사, 육아원, 양로원, 수녀원 등과 같은 시설단위의 조사구는 표본추출 대상 조사구에서 제외시키기로 하였다.

3) 基本標本の 規模

다음 번의 새로운 기본표본조사구는 1995년 인구총조사의 조사구를 이용하여 추출하여야 하므로, 1995년 인구총조사의 조사구를 이용할 수 있는 1996년 하반기에나 그 추출이 가능하게 된다. 따라서, 기본표본의 규모는 1993년부터 1996년도까지 4년간 사용할 수 있는 규모가 되어야 한다.

과거의 경험에 의하면 거의 모든 조사는 예산의 제약으로 100~150개 조사구를 조사하였으므로 그 최대치인 150개 조사구를 택하고, 2년에 3개 조사를 전제로 하면 기본표본조사구의 수는 1996년까지 총 900개의 조사구가 필요하게 된다.

이 때, 이 900개의 조사구가 전국에 골고루 분산되면 표본추출비용이 많이 소요되어 곤란하게 된다. 이에 대한 대책으로, 인접한 3개의 조사구를 통합하여 하나의 기본조사구로 하고, 300개의 기본조사구를 추출하는 방법이 있다.

그런데 인구총조사 조사구의 가구수는 40가구 내외가 되는 경우도 있으며, 이렇게 가구수가 적은 조사구를 실사표본조사구(실제 조사의 표본조사구)로 하는 것은 조사관리면에서 비효율적이므로, 이러한 조사구는 인접 조사구에 포함시켜서 2개의 조사구를 하나의 실사표본조사구로 취급하여야 한다. 따라서 이런 경우에는 4개의 조사구를 통합하여 하나의 기본조사구로 하여야 한다.

또한 어떤 조사에서는 실사표본조사구의 가구수가 100가구 이상이 되어야 하는 경우도 있을 수 있으며 약 200개의 조사구가 필요한 경우도 있을 수 있다.

이런 경우를 대비하여 기본표본조사구수는 300개가 아닌 400개로 하고, 각 기본조사구의 가구수는 240가구 이상이 되도록 하고 인구총조사의 조사구로는 적어도 4개 조사구가 되어야 한다.

나. 基本標本調査區의 抽出

1) 基礎資料의 整理

1990년 인구총조사구별 가구명부의 주택수 및 가구수가 수록된 전산화일에서 시설단위 조사구와 각 시·도의 섬지역의 조사구를 제외한 185,149개의 보통조사구로 작업화일을 작성하였다. 앞으로는 '인구총조사의 보통조사구'를 편의상 '인구총조사구'라는 용어로 바꾸어 사용하기로 한다.

2) 調査區의 層化 및 分類

인구총조사구를 행정구역번호를 기준으로 다음과 같이 3개 층으로 층화하였다.

- 제1층 : 6대 시(서울, 부산, 대구, 인천, 광주, 대전)
- 제2층 : 기타 시(각 도의 모든 시)
- 제3층 : 군부

이 때, 1990년 인구총조사 당시의 경기도 고양군은 고양시로 승격되었으므로, 경기도 고양군은 기타 시로 포함시켰다.

6대 시와 기타 시의 2개 층에서는 인구총조사구를 주택특성에 따라 단독주택이 가장 많은 조사구, 아파트가 가장 많은 조사구, 연립주택 및 다세대주택이 가장 많은 조사구, 기타 조사구의 순서로 분류하였으며, 각 분류 내에서는 다시 동·읍·면 단위까지의 행정구역번호 및 조사구번호 순서로 나열하였다.

군부의 경우에는 먼저 읍부와 면부의 순서로 분류한 다음에, 각각 위와 같은 방법과 순서로 나열하였다.

이와 같이 각 층별로 조사구를 나열한 다음, 각 조사구의 일반가구수를 10으로 나누어 반올림한 결과를 각 조사구의 크기의 測度(measure of size)로 부여하고, 이를 누적하였다.

3) 基本調査區의 設定

1개 기본표본조사구의 가구수가 240가구 이상이 되도록 하기 위하여 층별로 각각 처음 조사구부터 차례로 크기의 측도의 합계가 24를 초과할 때까지의 조사구를 통합하여 하나의 기본조사구로 설정하였다. 이 때, 기본조사구에 포함되는 인구총조사구의 수가 4개 미만인 경우는 그 다음의 조사구를 포함시켜 4개의 조사구로 하였다. 그리고 기본조사구의 크기의 측도는 인구총조사구의 크기의 측도의 합계를 부여하였다.

4) 層別 基本標本調査區數의 決定

기본표본조사구의 수를 400개 조사구로 하고, 층별 크기의 측도에 비례하도록 배분하면 <表 III-1>의 제1안과 같다.

<表 III-1> 層別 基本標本調査區數 配分內譯 - 既存 基本標本

지역	인구총조사		기본표본조사구수	
	조사구수	크기의 측도	제1안	제2안
6대 시	89,987 (48.60)	548,114 (48.26)	193 (48.26)	155 (38.75)
기타 시	50,077 (27.05)	310,747 (27.36)	109 (27.36)	88 (22.00)
군부	45,085 (24.35)	276,975 (24.38)	98 (24.38)	157 (39.25)
계	185,149(100.00)	1,135,836(100.00)	400(100.00)	400(100.00)

그런데 대부분의 조사에서는 시부와 군부의 별도 추정치가 요구되고 있으나, 100개 조사구를 조사하는 경우에는 군부의 조사구수는 24개 조사구 밖에 되지 않으며, 150개 조사구의 경우에는 37개 조사구 밖에 되지 않아, 군부의 조사구수가 기본적으로 적기 때문에 군부 추정치의 표본오차는 크게 되어 사실상 분석단계에서 군부에 대하여 언급하는 것은 별다른 의미를 찾을 수 없게 된다.

이와 같은 문제점을 조금이나마 해소하기 위하여, 군부의 크기의 측도에는 2배를 하여 배분하였으며, 그 결과는 <表 III-1>의 제2안과 같다. 즉, 시부:군부의 기본표본조사구의 구성비를 제1안은 크기의 측도에 비례하는 76:24가 되도

록 한 것이며, 제2안은 군부의 크기의 측도에는 2배의 가중치를 적용한 61 : 39가 되도록 한 것이다.

100개 조사구를 조사하는 경우에 군부의 조사구 수는 39개 조사구가 되며 150개 조사구의 경우에는 59개의 조사구가 되어, 제1안의 경우보다는 제2안의 경우가 보다 합리적이다. 따라서, 층별 기본표본조사구의 수는 제2안을 택하였다.

5) 基本標本調査區의 抽出

〈表 III-1〉의 제2안의 내용대로 6대 시에서 155개, 기타 시에서 88개, 군부에서 157개의 기본조사구를 크기의 측도에 비례하는 확률로 계통추출하여 기본표본조사구로 하였다. 이와 같이 추출한 결과의 표본규모는 〈表 III-2〉와 같다. 여기서, 기타 시를 별개의 층으로 하였다고 해서 기타 시에 대하여도 추정치를 산출하기 위한 것은 절대로 아니며, 단지 시부에서의 표본오차를 감소시키기 위한 하나의 방법일 뿐이다.

〈表 III-2〉 層別 基本標本 規模 - 既存 基本標本

지 역	기본표본 조사구수	인구총 조사구수	크기의 측도	가구수
6대 시	155 (38.75)	753 (36.37)	4,738 (38.09)	47,249 (38.21)
기타 시	88 (22.00)	427 (20.63)	2,729 (21.94)	27,134 (21.95)
군부	157 (39.25)	890 (43.00)	4,971 (39.97)	49,262 (39.84)
계	400(100.00)	2,070(100.00)	12,438(100.00)	123,645(100.00)

다. 實查標本調査區의 抽出方法

어떤 조사에서 주어진 예산과 조사목적에 따라 층별 표본조사구수가 결정되면, 층별 기본표본조사구 명부에서 정해진 수의 기본표본조사구를 먼저 계통추출한다. 그리고 표본으로 추출된 기본표본조사구에서는 하나의 인구총조사구를 크기의 측도에 비례하는 확률로 임의추출하여 실사표본조사구로 한다.

2. 改善方向 및 考慮事項

가. 基本標本の 規模

1993년도의 400개 기본표본조사구의 인구총조사구수는 총 2,070개 조사구이었으며, 1990년 인구총조사 결과의 가구수로는 123,645가구이었다. 따라서 기본표본조사구당 인구총조사 조사구수는 평균 5.2개 조사구가 되었으며, 기본표본조사구의 가구수는 최소 240가구인 평균 309가구이었다. 그런데 1년에 2개의 통계조사를 별도로 실시하고 그 다음 연도에 1개의 통계조사를 실시한 경우에는 조사구는 다르지만 전년도에 조사한 동·읍·면 내의 조사구가 다시 추출되므로, 해당 동·읍·면에서는 2년 연속하여 조사에 협조하게 되어 불평이 많았다.

따라서 1996년에는 이러한 불평을 해소하고 1997년부터 2001까지 5년간 사용하기 위하여 기본표본조사구수는 과거의 400개 조사구에서 200개 조사구를 증가시켜 600개 조사구로 하고 기본표본조사구의 최소 가구수는 과거보다 80가구가 많은 320가구를 넘도록 한다.

나. 層化

1993년도에는 인구총조사의 조사구를 (1) 6대 시, (2) 기타 시, (3) 군부의 3개 층으로 층화하였다. 그런데 1995년 인구총조사 당시에는 이미 광역시 및 도농복합시 등의 행정구역 개편으로 시에도 읍·면이 포함되어 있었다. 따라서 시부 및 군부별 추정외의 경우와 동부 및 읍·면부별 추정외의 경우를 대비하여 (1) 6대 시의 동, (2) 기타 시의 동, (3) 시의 읍·면, (4) 군의 읍, (5) 군의 면으로 5개 층을 설정한다.

다. 基本調査區의 設定

1993년도의 기본표본조사구를 추출할 때에는 인구총조사 조사구의 가구수를 10으로 나누어 반올림한 결과를 크기의 測度(measure of size)로 부여하고, 이 크기의 측도의 합계가 24이상이 되는 4~5개의 인구총조사구를 통합하여 하나

의 기본조사구를 설정한 다음에 400개의 기본조사구를 기본표본조사구로 추출하였다.

여기서 각 조사구에 크기의 측도를 부여한 이유는 조사구별 가구수는 항상 일정한 것이 아니고 가구의 전출입으로 어느 정도 변동이 있게 마련이므로 이러한 변동을 어느 정도까지는 당연한 것으로 인정하기 위한 것이다.

또한 표본으로 추출된 한 조사구의 가구수가 너무 적어 인접조사구에서 일부의 가구를 추가하는 경우의 편리성을 고려한 것이다. 즉 인접조사구에서 20가구 내외를 추가하고자 하는 경우 인접조사구의 크기의 측도가 6이면 가구수가 비슷하게 거처를 단위로 6개의 구역으로 분할하여 2개 구역만 추가하면 된다.

1996년도 기본표본조사구의 경우에는 기본조사구를 설정하지 않고 직접 인구총조사 조사구를 추출한 다음에, 표본으로 추출된 조사구부터 크기의 측도의 합계가 33이상이 될 때까지의 조사구를 하나의 기본표본조사구로 하였으며 조사구별 크기의 측도의 합계를 그 기본표본조사구의 크기의 측도로 하였다. 그리고 기본조사구를 설정하여 기본표본조사구를 추출한 것으로 간주한다.

이와 같이 기본조사구를 설정하지 않고도 기본조사구를 설정하여 기본표본조사구를 추출한 것으로 간주하는 편법을 사용한 이유는 1차추출단위 조사구를 설정하여 기본표본조사구를 추출한 결과와는 현실적으로 별다른 차이가 없을 뿐만 아니라, 1차추출단위 조사구를 설정하여 기본표본조사구를 추출하게 되면 절차가 복잡하고 노력도 많이 소요되기 때문이다.

라. 基本標本調査區의 分類

1993년도 기본표본조사구의 경우에는 조사가 있을 때마다 400개의 기본표본조사구 중에서 일차적으로 기본표본조사구를 계통추출하고 표본으로 추출된 기본표본조사구에서는 이미 사용하지 않은 인구총조사구를 실사표본조사구로 추출하였으나, 이를 보다 단순하게 하기 위하여 1996년도의 600개 기본표본조사구는 150, 200 및 250개 조사구의 3종으로 나누어 150개 기본표본조사구를 제1표본, 200개 기본표본조사구를 제2표본, 250개 기본표본조사구를 제3표본으로 한다.

따라서 실사표본조사구가 150개 이하인 경우에는 제1표본을 사용하고 151~

200개 조사구가 되는 경우에는 제2표본을, 그리고 201~250개 조사구가 되는 경우에는 제3표본을 사용토록 한다.

첫 번째 조사에서 제1표본을 사용하였는데, 다음 번의 조사에서도 실사표본조사구가 150개 조사구 이하인 경우에는 제1표본을 사용하지 않고 제2표본을 사용한다. 마찬가지로 어떤 조사에서 제2표본을 사용하였는데 다음 번의 다른 조사에서도 실사표본조사구가 150~200개 조사구인 경우에는 제3표본을 사용한다. 이러한 방법을 택하면, 동일한 동·읍·면이 2번 연속하여 해당되는 것을 방지하는데 편리하게 된다.

만일 어떤 특수한 조사에서 제3표본의 250개 기본표본조사구보다 많은 수의 실사표본조사구가 필요하게 된다면 제1표본, 제2표본, 제3표본의 3개 표본 중에서 2개의 표본을 통합하여 사용하면 된다.

현재까지 1개 조사에서 200개 조사구 이상을 조사한 적이 없었지만 2000년도 장애인실태조사의 경우에는 200개 조사구를 초과할 수도 있으므로 250개 조사구를 최대로 한 것이다.

IV. 基本標本調査區의 改編

여기서는 1996년 기본표본조사구의 개편내용에 관하여 설명하였다. 따라서 2001년의 기본표본조사구 개편에서는 여기에 설명된 절차와 방법을 그대로 따르거나 개선시키면 된다.

1. 基本標本調査區의 抽出

가. 基礎資料의 整理

1995년 인구주택총조사의 잠정집계에서 사용된 조사구별 자료에서 각 시·도의 섬지역 조사구와 시설단위 조사구를 제외한 219,771개 보통조사구 자료로 전산작업화일을 작성하였다. 제외된 섬지역의 가구수는 실제적으로 전국 가구수의 약 0.6% 미만에 불과하다.

앞으로는 '인구주택총조사의 보통조사구'를 편의상 '인구총조사구'라는 용어로 바꾸어 사용하기로 한다.

나. 調査區의 層化

인구총조사구를 행정구역번호를 기준으로 다음과 같이 5개 층으로 층화하였으며, 각 인구총조사구에 층번호를 부여하였다.

제1층 : 6대 시(서울, 부산, 대구, 인천, 광주, 대전)의 동

제2층 : 기타 시(각 도의 모든 시)의 동

제3층 : 시(6대 시 및 기타 시)의 읍·면

제4층 : 군의 읍

제5층 : 군의 면

이 때, 1995년 인구총조사 이후 1996년 7월말까지 행정구역이 개편된 지역의 인구총조사구는 신행정구역에 따라 층번호를 수정하였다.

다. 層別 人口總調査區의 整列

이와 같이 인구총조사구를 5개의 층으로 층화한 다음에는 제3층(시의 읍·면)을 제외한 4개 층에서는 각각 다음과 같이 인구총조사구를 정렬하였다.

1차 정렬 : 주택특성에 의한 정렬

- (1) 단독주택이 가장 많은 조사구
- (2) 아파트가 가장 많은 조사구
- (3) 연립 및 다세대주택이 가장 많은 조사구
- (4) 기타 조사구

2차 정렬 : 행정구역번호 및 조사구번호에 의한 정렬

- (1) 동·읍·면 단위까지의 행정구역번호
- (2) 조사구번호

그리고 제3층(시의 읍·면)의 경우에는 다음과 같이 인구총조사구를 정렬하였다.

1차 정렬 : 행정구역에 의한 정렬

- (1) 6대 시의 읍
- (2) 기타 시의 읍
- (3) 6대 시의 면
- (4) 기타 시의 면

2차 정렬 : 주택특성에 의한 정렬(다른 층의 제1차 정렬과 같음)

3차 정렬 : 행정구역번호 및 조사구번호에 의한 정렬(다른 층의 제2차 정렬과 같음)

라. 크기의 測度 賦與

각 층에서 인구총조사구를 위와 같이 나열한 다음에는 층별로 인구총조사구의 일반가구수를 10으로 나누어 반올림한 결과를 크기의 측도로 부여하였다. 그리고 확률비례계통추출법을 사용하기 위하여 층별로 이 크기의 측도를 누적하

였으며, 층별 누적결과는 <表 IV-1>과 같다.

<表 IV-1> 層別 크기의 測度

층	인구총조사구수	가구수	크기의 측도
(1) 6대 시의 동	104,222 (47.4)	6,035,168 (46.8)	608,395 (46.83)
(2) 기타 시의 동	68,589 (31.2)	4,051,452 (31.5)	408,580 (31.45)
(3) 시의 읍·면	18,825 (8.6)	1,120,502 (8.7)	112,966 (8.69)
(4) 군의 읍	10,220 (4.6)	606,878 (4.7)	61,178 (4.71)
(5) 군의 면	17,915 (8.2)	1,071,650 (8.3)	108,048 (8.32)
합 계	219,771(100.0)	12,885,650(100.0)	1,299,167(100.00)

이 表에서 각 층의 가구수를 10으로 나눈 결과와 크기의 측도가 비슷하지 않은 것은 가구수의 끝단위가 5~9인 조사구가 우연적으로 많기 때문이다.

마. 層別 基本標本調査區數의 決定

기본표본조사구의 수를 600개 조사구로 하고, 층별 크기의 측도에 비례하도록 배분하면 <表 IV-2>의 제1안과 같다.

<表 IV-2> 層別 基本標本調査區數 配分內譯

층	인구총조사		기본표본조사구수	
	조사구수	크기의 측도	제1안	제2안
(1) 6대 시의 동	104,222	608,395 (46.83)	281 (46.8)	231 (38.5)
(2) 기타 시의 동	68,589	408,580 (31.45)	189 (31.5)	155 (25.8)
(3) 시의 읍·면	18,825	112,966 (8.69)	52 (8.7)	86 (14.3)
(4) 군의 읍	10,220	61,178 (4.71)	28 (4.7)	46 (7.7)
(5) 군의 면	17,915	108,048 (8.32)	50 (8.3)	82 (13.7)
합계	219,771	1,299,167(100.00)	600(100.0)	600(100.0)

제1안에서 제1층의 경우에 기본표본조사구수는 다음과 같이 계산된다.

$$608,395 / 1,299,167 \times 600 = 0.4683 \times 600 = 46.83 \times 6 = 280.90 = 281$$

따라서 제2층~제5층의 기본표본조사구수는 다음과 같이 된다.

$$\text{제2층 : } 31.45 \times 6 = 155$$

$$\text{제3층 : } 8.69 \times 6 = 52$$

$$\text{제4층 : } 4.71 \times 6 = 28$$

$$\text{제5층 : } 8.32 \times 6 = 50$$

그런데 대부분의 조사에서는 동부와 읍·면부의 별도 추정치가 요구될 것이다. 100개 조사구를 조사하는 경우에 읍·면부의 조사구수는 8.7(3층) + 4.7(4층) + 8.3(5층) = 21.7 = 22개 조사구밖에 되지 않으며, 150개 조사구의 경우에는 21.7 + (21.7 / 2) = 33개 조사구, 200개 조사구의 경우에는 21.7 × 2 = 43개 조사구밖에 되지 않아 기본적으로 조사구수가 적기 때문에 읍·면부 추정치의 표본오차가 크게되어, 분석단계에서 읍·면부에 대하여 언급하는 것은 별다른 의미를 찾을 수 없게 된다(시부와 군부로 분리하는 경우에 군부의 조사구수는 읍·면부의 경우보다 더 적어진다).

이와 같은 문제점을 조금이나마 해소하기 위하여, 읍·면지역인 제3층~제5층에는 2배의 가중치를 적용하여 배분하였으며, 그 결과는 <表 IV-2>의 제2안과 같다.

즉, 제1층과 제2층의 크기의 측도는 그대로 더하고 제3층~제5층의 크기의 측도는 2배를 하여 더하면 합계는 다음과 같이 된다.

$$\begin{aligned} & 1\text{층} + 2\text{층} + (3\text{층} \times 2) + (4\text{층} \times 2) + (5\text{층} \times 2) \\ & = 608,395 + 408,580 + (112,966 \times 2) + (61,178 \times 2) + (108,048 \times 2) \\ & = 1,581,359 \end{aligned}$$

따라서 각층의 기본표본조사구수는 다음과 같이 계산된다.

$$\text{제1층 : } (608,395 \times 1) / 1,581,359 \times 600 = 230.84 = 231$$

$$\text{제2층 : } (408,580 \times 1) / 1,581,359 \times 600 = 155.02 = 155$$

$$\text{제3층} : (112,966 \times 2) / 1,581,359 \times 600 = 85.72 = 86$$

$$\text{제4층} : (61,178 \times 2) / 1,581,359 \times 600 = 46.42 = 46$$

$$\text{제5층} : (108,048 \times 2) / 1,581,359 \times 600 = 81.99 = 82$$

크기의 측도의 구성비를 이용하는 경우에는 다음과 같이 된다.

$$\begin{aligned} & 1\text{층} + 2\text{층} + (3\text{층} \times 2) + (4\text{층} \times 2) + (5\text{층} \times 2) \\ & = 46.83 + 31.45 + (8.69 \times 2) + (4.71 \times 2) + (8.32 \times 2) = 121.72 \end{aligned}$$

따라서 각층의 기본표본조사구수는 다음과 같이 계산된다.

$$\text{제1층} : (46.83 \times 1) / 121.72 \times 600 = 230.84 = 231$$

$$\text{제2층} : (31.45 \times 1) / 121.72 \times 600 = 155.03 = 155$$

$$\text{제3층} : (8.69 \times 2) / 121.72 \times 600 = 85.67 = 86$$

$$\text{제4층} : (4.71 \times 2) / 121.72 \times 600 = 46.43 = 46$$

$$\text{제5층} : (81.32 \times 2) / 121.72 \times 600 = 82.02 = 82$$

이와 같이 층별 크기의 측도를 이용하는 경우와 층별 크기의 측도의 구성비를 이용하는 경우의 층별 기본표본조사구수는 같게 마련이다.

제2안에서 100개 조사구를 조사하는 경우에 읍·면부의 조사구수는 14.3(3층) + 7.7(4층) + 13.7(5층) = 35.7 = 36개 조사구가 되며 150개 조사구의 경우에는 35.7 + (35.7 / 2) = 53.6 = 54개 조사구, 200개 조사구의 경우에는 35.7 × 2 = 71.4 = 71개 조사구가 되어 제1안의 경우보다는 제2안의 경우가 보다 합리적이다. 따라서, 층별 기본표본조사구의 수는 제2안을 택하였다.

바. 基本標本調査區의 抽出

<表 IV-2>의 제2안의 내용대로 6대 시의 동에서 231개, 기타 시의 동에서 155개, 시의 읍·면에서 86개, 군의 읍에서 46개, 군의 면에서 82개의 인구총조사구를 크기의 측도에 비례하는 확률로 계통추출하였다.

그리고 추출된 인구총조사구부터 크기의 측도의 합계가 32를 초과할 때(33이상)이 될 때까지 그 다음의 인구총조사구를 통합하여 기본표본조사구로 하였다.

이 때, 크기의 측도의 합계가 32를 초과할 때까지 크기의 측도를 더하는 과정에서 인구총조사의 10% 표본조사구가 있을 때에는 이 조사구의 크기의 측도는 더하지 않았다. 즉 인구총조사의 10% 표본조사구는 대상에서 제외하였다.

인구총조사의 10% 표본조사구를 제외한 이유는 통계청에서 1997년도에 이 10% 표본조사구를 이용하여 가구단위 통계조사(경제활동인구조사, 도시가계조사 등)의 표본개편을 실시하기 때문에 통계청의 표본조사구와 중복을 피하기 위한 것이다.

이와 같이 기본표본조사구를 추출한 결과는 <表 IV-3>과 같이 요약된다.

<表 IV-3> 基本標本調査區 抽出結果

층	기본표본 조사구수	인구총조사 조사구수	크기의 측도	가구수
(1) 6대 시의 동	231 (38.50)	1,323 (36.37)	8,137 (38.09)	81,168 (38.21)
(2) 기타 시의 동	155 (25.83)	863 (20.63)	5,481 (21.94)	54,602 (21.95)
(3) 시의 읍·면	86 (14.33)	479 (36.37)	3,032 (38.09)	30,160 (38.21)
(4) 군의 읍	46 (7.67)	239 (20.63)	1,632 (21.94)	16,185 (21.95)
(5) 군의 면	82 (13.67)	459 (43.00)	2,889 (39.97)	28,693 (39.84)
계	600(100.00)	3,363(100.00)	21,171(100.00)	210,808(100.00)

사. 基本標本調査區 名簿作成

600개의 기본표본조사구와 600개의 기본표본조사구에 포함된 3,363개의 인구총조사구의 명부는 <樣式 1>과 같이 작성하였다. 이 양식에서 각 항목의 내용은 다음과 같다.

① 행정구역명

- 1996. 6. 30 현재의 행정구역명이다.
- 읍·면의 경우에는 행정 및 법정상의 읍·면이 동일하지만, 동의 경우에는 행정동과 법정동이 같은 경우도 있고 다른 경우도 있다.
- 1995년 인구총조사 당시의 이에 관한 사례는 다음과 같다.

〈樣式 1〉 基本標本調査區 名簿

행정구역명	행정구역 번호	조사구 번호	가구 수	크기 측도	추출 순서	일련 번호	특성 번호
①	②	③	④	⑤	⑥	⑦	⑧
서울 용산구 효자동	11-030-58	032	69	7	1-192	1084	3
	11-030-58	042	40	4	1-192	1085	3
	11-030-58	051	66	7	1-192	1086	3
	11-030-58	052	73	7	1-192	1087	3
	11-030-58	053	64	6	1-192	1088	3
	11-030-58	054	45	5	1-192	1089	3
서울 광진구 구의4동	11-050-62	039	64	6	1-193	1090	3
	11-050-62	040	72	7	1-193	1091	3
	11-050-62	041	59	6	1-193	1092	3
	11-050-62	042	82	8	1-193	1093	3
	11-050-62	043	64	6	1-193	1094	3

- 서울특별시 용산구 효자동사무소의 관할구역은 법성동으로 효자동뿐이다(행정동과 법정동이 같은 경우).
- 인천광역시 중구 내경동사무소의 관할구역은 내동, 경동, 용동의 3개 법정동이다(3개의 법정동이 1개의 행정동인 경우).
- 광주광역시 동구 산수1동사무소, 산수2동사무소, 산수3동사무소의 관할구역은 법정동인 산수동의 일부씩이다(1개의 법정동이 3개의 행정동으로 분리된 경우).

② 행정구역번호

- 1995년 인구총조사의 자료처리에서 사용된 행정구역번호이다.
- 처음의 2단위는 시·도번호이며, 다음의 3단위는 구·시·군번호이다. 그리고 마지막 2단위는 행정구역상의 동·읍·면 번호이다.

③ 조사구번호

- 1995년 인구총조사구 번호이다. 조사구 번호는 동·읍·면별로 001번부터

일련번호를 부여하였다.

- 같은 동·읍·면 내에서는 조사구가 서로 인접하도록 조사구번호가 부여되어 있다.

④ 가구수

- 1995년 인구총조사 결과의 일반가구수이다.

⑤ 크기의 측도

- 가구수 10으로 나누어 반올림한 것이다.

⑥ 추출순서

- 처음의 1단위는 층번호이며, 다음의 3단위는 그 층에서 기본표본조사구가 추출된 순서번호이다.

⑦ 일련번호

- 기본표본조사구의 추출순서 및 조사구번호(③) 순서에 의하여 부여한 일련번호이다.
- 마지막 일련번호는 기본표본조사구에 포함된 인구총조사구수와 같게 된다.

⑧ 특성번호

인구총조사구의 주택특성번호를 말한다.

- 1 : 단독주택이 가장 많은 조사구
- 2 : 아파트가 가장 많은 조사구
- 3 : 연립 및 다세대주택이 가장 많은 조사구
- 4 : 기타 조사구

〈樣式 1〉의 기본표본조사구명부에서 서울 용산구 효자동의 기본표본조사구에 포함된 6개 인구총조사구 번호는 032, 042, 051~054이므로 여기에는 결번 조사구가 있다.

결번 조사구는 각 층에서 인구총조사구를 주택특성과 행정구역번호 및 조사구번호 순서로 나열하였기 때문에 이 기본표본조사구에 포함된 인구총조사구의 특성번호는 3(연립 및 다세대주택이 가장 많은 조사구)이지만, 결번조사구는 특성번호가 3이 아닌 경우이거나 특성번호가 4이지만 크기의 측도를 누적하는 과정에서 10%의 표본조사구이기 때문에 제외된 조사구도 있을 수 있다.

아. 調查區要圖의 複寫

600개 기본표본조사구에 포함된 3,363개의 인구총조사 조사구의 조사구요도를 각각 2부씩 복사하였다. 이 조사구요도는 조사구별로 조사구 경계선, 주요 지형지물, 거처 등이 표시되고 거처번호가 기입된 약식 지도로 1995년 인구총조사의 조사단계에서 동·읍·면 직원과 조사원이 수정·보완한 것이다.

조사구요도의 기본자료로는 통계청에서 제공한 각종 지도가 사용되었는데 지역에 따라 그 종류가 다르며, 그 종류는 다음과 같다.

① 基本圖

지형도, 항공사진측량(약식)현황도, 상하수도 배관망도 등의 조사구설정에 적합한 지도를 전산입력하여 동·읍·면 단위별로 접합, 절단한 후 지역특성을 고려한 축척배율로 확대·축소하여 조사구 설정용으로 제작한 지도

㉠ 국가 기본지형도(축척 : 1/5,000)

- 국립지리원에서 제작한 지형도 중에서 규격이 일정하고 정확도가 통일된 것으로 축척이 최대인 지형도
- 산간지역 및 군사보호지역 등 특정지역 이외에 전국이 거의 제작됨

㉡ 항공사진측량(약식) 현황도(축척 : 1/500 ~ 1/3,000)

- 지방행정기관이 항공촬영하여 자체 제작한 지도
- 축척은 1/500, 1/600, 1/1,500, 1/2,500, 1/3,000 등으로 다양
- 사용지역 : 서울, 부산, 대구, 인천, 광주, 대전, 수원, 성남, 안양, 부천, 강릉, 춘천, 김제, 안산, 청주, 속초, 삼척, 전주, 마산, 구리, 군포 등 21개 지역

㉢ 국립지리원 1/25,000 지형도

- 규격이 일정하고 축척을 작게하여 제작한 지형도
- 사용지역 : 1/5,000 기본지형도가 제작되지 않은 산간지역

② 擴大圖

기본도 상에서 거처가 밀집되어 거처별 식별이 곤란한 지역은 조사구 설

정이 가능하도록 별도로 확대하여 제작한 지도(부분확대와 전체확대)

전산입력된 각종 지도에는 동·읍·면별로 설정한 조사구의 경계선도 입력되어 있으며, 통계청에서는 각 조사구마다 A3 크기로 전산출력하여 조사구요도로 사용토록 하였다. 따라서 각 동·읍·면에서는 인구총조사의 실사과정에서 조사구별로 전산출력된 조사구요도를 수정·보완하고, 인구총조사의 거쳐번호를 기입하였다.

기본표본조사구 추출과 관련된 기획단계에서는 동·읍·면에 배부되었던 조사구요도를 전산출력할 계획이었으나, 조사구요도를 전산출력하는 것보다는 동·읍·면사무소에서 수정·보완하고 거쳐번호를 기입한 조사구요도를 복사하는 것이 현지조사에서 더 쓸모가 있으므로 수정·보완된 조사구요도를 복사하였다.

2. 基本標本調査區의 編制

가. 編制의 內容

600개의 기본표본조사구를 150개의 조사구(제1표본), 200개의 조사구(제2표본), 250개의 조사구(제3표본)의 3종류로 분류하기 위하여, 표본의 종류별로 각각 기본표본조사구를 600개 전체 기본표본조사구의 층별 구성비에 비례하도록 각층에 배분하였으며, 그 결과는 <表 IV-4>와 같다.

<表 IV-4> 標本の 種類別 基本標本調査區의 層別 配分內譯

층	기본표본조사구	제1표본	제2표본	제3표본
(1) 6대 시의 동	231 (38.50)	58 (38.67)	77 (38.50)	96 (38.40)
(2) 기타 시의 동	155 (25.83)	38 (25.33)	52 (26.00)	65 (26.00)
(3) 6대 시의 읍·면	86 (14.33)	21 (14.00)	29 (14.50)	36 (14.40)
(4) 군의 읍	46 (7.67)	12 (8.00)	15 (7.50)	19 (7.60)
(5) 군의 면	82 (13.67)	21 (14.00)	27 (13.50)	34 (13.60)
계	600(100.00)	150(100.00)	200(100.00)	250(100.00)

이 表에서 제2표본의 200개 기본표본조사구의 경우에 각층의 기본표본조사구 수 산출방법을 예로 들면 다음과 같다.

$$\text{제1층} : 200 \times 38.50 / 100 = 77.00 = 77$$

$$\text{제2층} : 200 \times 25.83 / 100 = 51.66 = 52$$

$$\text{제3층} : 200 \times 14.33 / 100 = 28.66 = 29$$

$$\text{제4층} : 200 \times 7.67 / 100 = 15.34 = 15$$

$$\text{제5층} : 200 \times 13.67 / 100 = 27.34 = 27$$

나. 編制의 方法

600개의 기본표본조사구를 <表 IV-4>의 제1표본, 제2표본 및 제3표본의 층별 기본표본조사구수에 따라 우선 150개의 기본표본조사구를 먼저 추출하여 제1표본으로 하였으며, 나머지의 450개 기본표본조사구 중에서 200개의 기본표본조사구를 추출하여 제2표본으로 하였다.

따라서 나머지의 250개 기본표본조사구를 제3표본이 되었으며, 제1표본과 제2표본의 실제적인 추출절차 및 방법은 다음과 같다.

1) 第1標本

600개 기본표본조사구 중에서 제1표본의 150개 기본표본조사구는 다음과 같이 추출하였다.

- (1) 제1층의 231개의 기본표본조사구 중에서 58개의 기본표본조사구를 계통 추출하였다.
- (2) 제2층의 155개의 기본표본조사구 중에서 38개의 기본표본조사구를 계통 추출하였다.
- (3) 제3층의 86개의 기본표본조사구 중에서 21개의 기본표본조사구를 계통추출하였다.
- (4) 제4층의 46개의 기본표본조사구 중에서 12개의 기본표본조사구를 계통추출하였다.

- (5) 제5층의 82개의 기본표본조사구 중에서 21개의 기본표본조사구를 계통추출하였다.

2) 第2標本

600개 기본표본조사구 중에서 제1표본을 추출하고 남은 450개 기본표본조사구 중에서 제2표본의 200개 기본표본조사구는 다음과 같이 추출하였다. 그리고 나머지의 기본표본조사구를 제3표본으로 하였다.

- (1) 제1층의 나머지 173개 기본표본조사구 중에서 77개의 기본표본조사구를 계통추출하였다.
- (2) 제2층의 나머지 117개 기본표본조사구 중에서 52개의 기본표본조사구를 계통추출하였다.
- (3) 제3층의 나머지 65개 기본표본조사구 중에서 29개의 기본표본조사구를 계통추출하였다.
- (4) 제4층의 나머지 34개 기본표본조사구 중에서 15개의 기본표본조사구를 계통추출하였다.
- (5) 제5층의 나머지 61개 기본표본조사구 중에서 27개의 기본표본조사구를 계통추출하였다.

다. 標本의 種類別 基本標本調査區 名簿

기본표본조사구 명부는 제1표본, 제2표본, 제3표본에 대하여 별도로 작성하였으며, 제2표본의 경우를 예로 들면 <樣式 2>와 같다.

(1) 조사구번호

① 좌측 2단위의 첫 번째 수는 표본의 종류번호이고 두 번째 수는 층번호이다.

- 표본의 종류번호는 다음과 같다.

- 1 : 제1표본 (150개 기본표본조사구)
- 2 : 제2표본 (200개 기본표본조사구)
- 3 : 제3표본 (250개 기본표본조사구)

〈樣式 2〉 標本의 種類別 基本標本調査區 名簿

조사구 번호	행정구역	인구총조사				일련 번호	비고
		가구 수	특성 번호	행정구역 번호	조사구 번호		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
23-09-1	경북 영주시 풍기읍	49	4	37-060-11	077	0782	
23-09-2		48	4	37-060-11	078	0783	
23-09-3	경북 문경시 문경읍	56	4	37-090-11	019	0784	
23-09-4		58	4	37-090-11	022	0785	
23-09-5		61	4	37-090-11	024	0786	
23-09-6		68	4	37-090-11	025	0786	

- 층번호는 다음과 같다.

- 1 : 6대 시의 동
- 2 : 기타 시의 동
- 3 : 시의 읍면
- 4 : 군의 읍
- 5 : 군의 면

- 예를 들면 23-08-1의 경우 첫 번째 2단위 23은 제2표본의 3층임을 나타낸다.

② 두 번째의 2단위는 기본표본조사구가 추출된 순서번호이다.

- 예를 들면, 23-08-1의 경우 두 번째 2단위 08은 해당 층(제2표본의 3층)에서 8번째로 추출된 기본표본조사구임을 나타낸다.

③ 우측의 1단위는 기본표본조사구 내의 인구총조사구에 일련번호를 부여한 것이다.

(2) 행정구역

- 1996. 6. 30 현재의 동·읍·면단위의 행정구역이다.

(3) 가구수

- 1995년 인구총조사 당시의 일반가구수이다.

(4) 특성번호

- 1995년 인구총조사구의 주된 주택의 종류번호를 나타낸다.
 - 1 : 단독주택이 가장 많은 조사구.
 - 2 : 아파트가 가장 많은 조사구
 - 3 : 연립 및 다세대주택이 가장 많은 조사구
 - 4 : 기타 조사구

(5) 행정구역번호

- 1995년 인구총조사 당시의 행정구역번호이다.
- 처음의 2단위는 시도 번호이며, 두 번째의 3단위는 구·시·군 번호이다. 그리고 세 번째의 2단위는 동·읍·면 번호이다.

(6) 조사구번호

- 1995년 인구총조사의 조사구번호이다.

(7) 일련번호

- ①란의 조사구번호 순서에 의한 인구총조사구의 일련번호이다.

(8) 비고

- 사용된 조사구의 경우, 그 조사의 번호를 기입하는 란으로 현재는 공란이다.
- 예를 들면, 1997년도 전국 출산력 및 가족보건실태조사에 사용된 조사구의 경우에는 1을 기입한다.

〈樣式 2〉에 제시된 23-09-1부터 6까지의 기본표본조사구를 보면 인구총조사의 조사구는 경북 영주시 풍기읍의 2개 조사구와 문경시 문경읍의 4개 조사구가 포함되어 있다. 이와 같이 2개 시의 인구총조사구가 하나의 기본표본조사구에 포함되는 경우는 극히 드물지만, 각 층에서 인구총조사구의 주택특성과 행정구역번호 및 조사구번호 순서로 정렬하여 기본표본조사구를 추출하였기 때문에 어쩔수 없이 발생한 것이다.

V. 實查標本調査區의 抽出方法

여기서는 제1표본, 제2표본 및 제3표본에서 실사표본조사구를 추출하는 방법을 설명하였으며, 표본추정치 산출공식에는 실사표본조사구가 표본으로 추출된 확률이 고려되므로 실사표본추출과 관련하여 표본추출확률의 계산방법에 대하여도 설명하였다.

1. 基本的인 考慮事項

실사표본조사구수가 150개 조사구인 경우에는 제1표본을, 200개 조사구인 경우에는 제2표본을 사용하고 250개 조사구인 경우에는 제3표본을 사용하여 각 기본표본조사구에서 1개씩의 인구총조사구를 크기의 측도에 의한 확률비례추출법으로 추출하여 실사표본조사구로 한다.

실사표본조사구수가 150개 조사구, 200개 조사구, 250개 조사구가 아닌 경우에는 제1표본, 제2표본, 제3표본의 기본표본조사구 중에서 일부의 기본표본조사구를 계통추출하고, 표본으로 추출된 기본표본조사구에서 1개씩의 인구총조사구를 크기의 측도에 의한 확률비례계통추출법으로 추출하여 실사표본조사구로 한다.

그런데 동일한 동·읍·면에 2년 연속하여 조사협조를 의뢰하는 것은 피하는 것이 좋다. 그러기 위해서는 어떤 조사에서 제2표본을 사용하였다면 다음의 조사에서는 제1표본이나 제3표본을 사용하고, 제1표본을 사용하였다면 다음 번 조사에서는 제2표본이나 제3표본을 사용하여 한다.

2000년도의 장애인실태조사의 경우에는 실사표본조사구수가 200개 이상이 되어야 하고 각 실사표본조사구의 가구수는 200가구 이상이 되어야 하므로, 제3표본에서 그 때 사용할 인구총조사구에는 기본표본조사구 명부의 비고란에 'D'라고 기입하여 놓았으므로 이들 조사구는 제외하고 사용하여야 한다.

2. 抽出方法

가. 方法 (1)

1) 實查標本調査區의 抽出

제1표본, 제2표본, 제3표본의 각 기본표본조사구에서 1개씩의 인구총조사구를 실사표본조사구로 추출할 때에는 크기의 측도에 비례하는 확률로 임의추출한다. 제2표본의 3층에서 8번째의 기본표본조사구를 예로 들면, 이 기본표본조사구는 다음과 같이 5개의 인구총조사구로 구성되어 있다.

조사구 번호	행정구역	인 구 총 조 사				일련 번호	비고
		가구 수	특성 번호	행정구역 번호	조사구 번호		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
23-08-1	경북 경산시 하양읍	60	2	37-100-11	041	0776	
23-08-2		68	2	37-100-11	042	0777	
23-08-3		56	2	37-100-11	043	0778	
23-08-4		69	2	37-100-11	044	0779	
23-08-5		67	2	37-100-11	045	0780	

다음 작업표는 5개 인구총조사구의 가구수를 10으로 나누어 반올림하여 크기의 측도를 부여하고 이 크기의 측도를 누적한 결과가 33임을 보여준다.

조사구 번호	가구수	크기의 측도	누적결과	표본지정번호
23-08-1	60	6	6	-
23-08-2	68	7	13	-
23-08-3	56	6	19	17
23-08-4	69	7	26	-
23-08-5	67	7	33	-

따라서 난수표를 이용하여 33이하의 수를 택한다. 여기서 17이 택해졌다고 하면, 누적결과에서 17이 포함되는 조사구는 '23-08-3'번 조사구이므로 이 조사구를 실사표본조사구로 택한다.

만일 어떤 인구총조사구가 이미 다른 조사의 실사표본조사구로 사용되었다면 이 조사구는 제외하고 나머지 조사구 중에서 추출한다. 앞의 예에서 3번째 조사구가 이미 다른 조사에 사용되었다면, 이 조사구는 제외하고 크기의 측도를 누적하여 4개의 조사구 중에서 1개의 조사구를 크기의 측도에 비례하는 확률로 임의추출한다. 예를 들면 다음의 작업표는 크기의 측도의 최종 누적결과가 27이므로 27이하의 난수 하나를 택한 결과 15가 선택된 경우이다.

조사구 번호	가구수	크기의 측도	누적결과	표본지정번호
23-08-1	60	6	6	-
23-08-2	68	7	13	-
23-08-3	56	비해당	-	-
23-08-4	69	7	20	15
23-08-5	67	7	27	-

2) 實查標本調査區의 抽出確率 算出公式

제1표본, 제2표본 또는 제3표본의 각 기본표본조사구에서 1개씩의 인구총조사구를 실사표본조사구로 추출하는 경우에 각 실사표본조사구가 전체 인구총조사구 중에서 표본으로 추출되는 확률은 다음의 3개 확률을 곱하면 된다.

- ① P_1 = 각 층에서 기본표본조사구가 추출된 확률
- ② P_2 = 각 층의 기본표본조사구 중에서 제1표본, 제2표본, 제3표본의 3개 표본으로 추출된 확률
- ③ P_3 = 각 기본표본조사구 중에서 인구총조사구가 추출된 확률

앞으로 모든 공식에서는 다음과 같은 기호를 사용하기로 한다.

P = 추출확률

n = 기본표본조사구수

S = 크기의 측도

g = 제g표본을 나타내는 첨자

h = h번째 층을 나타내는 첨자

i = h번째 층의 i번째 기본표본조사구를 나타내는 첨자

α = 제g표본의 α 번째 기본표본조사구를 나타내는 첨자

j = 제g표본에서 α 번째 기본표본조사구 내의 j번째 인구총조사구를 나타내는 첨자

공식 : 실사표본조사구 추출확률 (1)

$$\begin{aligned}
 {}_gP_{h\alpha j} &= P_{\text{①}} \times P_{\text{②}} \times P_{\text{③}} \\
 &= \frac{n_h S_{hi}}{S_h} \cdot \frac{{}_g n_h}{n_h} \cdot \frac{{}_g S_{h\alpha j}}{{}_g S_{h\alpha}} \quad : \quad S_{hi} = {}_g S_{h\alpha} \\
 &= \frac{{}_g n_h}{S_h} {}_g S_{h\alpha j} \quad \dots\dots (5.1)
 \end{aligned}$$

이 공식에서 사용된 기호를 알기 쉽게 다른 방법으로 설명하면 다음과 같다.

${}_gP_{h\alpha j}$ = 제g표본의 h번째 층에서 α 번째 기본표본조사구 내의 j번째 인구총조사 조사구의 추출확률

n_h = h번째 층의 기본표본조사구수

${}_g n_h$ = 제g표본에서 h번째 층의 기본표본조사구수

S_h = h번째 층의 크기의 측도(表 IV-1 참조)

S_{hi} = h번째 층의 i번째 기본표본조사구의 크기의 측도

${}_g S_{h\alpha}$ = 제g표본의 h번째 층에서 α 번째 기본표본조사구의 크기의 측도

${}_g S_{h\alpha j}$ = 제g표본의 h번째 층에서 α 번째 기본표본조사구 내의 j번째 인구총조사구의 크기의 측도

제2표본의 제3층(시의 읍·면)의 29개 기본표본조사구 중에서 8번째 기본표본조사구 내의 3번째의 인구총조사구가 크기의 측도에 비례하는 확률로 임의추출된 경우를 예로 든다. 이 조사구의 인구총조사 당시의 일반가구수는 69가구이다.

공식 (5.1)의 각 항목과 관계되는 값을 정리한 결과는 다음과 같다.

$$\begin{aligned} {}_g n_h &= {}_2 n_3 &&= 29 \text{ (表 IV-4 참조)} \\ {}_g S_{h\alpha\beta} &= {}_2 S_{3, 08, 4} &&= 7(69\text{가구를 } 10\text{으로 나누어 반올림}) \\ S_h &= S_3 &&= 112,966 \text{ (表 IV-1 참조)} \end{aligned}$$

여기서 ${}_2 S_{3, 08, 4}$ 의 우측 첨자 3, 08, 4는 $h=3$, $\alpha=08$, $\beta=4$ 를 3084로 표시하면 구분이 되지 않으므로, 이를 구분한 것이다.

위의 값을 공식 (5.1)에 대입하여 계산한 결과는 다음과 같다.

$$\begin{aligned} {}_g P_{h\alpha j} &= \frac{{}_g n_h}{S_h} {}_g S_{h\alpha j} \\ {}_2 P_{3, 08, 4} &= \frac{29}{112,966} \cdot 7 = \frac{1}{3,895.3793} \cdot 7 \\ &= \frac{1}{556.4828} \end{aligned}$$

따라서 제2표본의 제3층에서 8번째 기본표본조사구내의 4번째 인구총조사구가 실사표본조사구로 추출된 확률은 $1/556.48$ 이 된다.

나. 方法 (2)

실사표본조사구수가 150개 조사구인 경우에는 제1표본을, 200개 조사구인 경우에는 제2표본을 사용하고 250개 조사구인 경우에는 제3표본을 사용하여 각 기본표본조사구에서 1개씩의 인구총조사구를 추출하면 되지만, 실사표본조사구수가 150개, 200개 또는 250개 조사구가 아닌 경우도 있게 된다.

예를 들어 158개의 실사표본조사구가 필요하다면 제2표본에서 우선 158개의 기본표본조사구를 추출하고 표본으로 추출된 기본표본조사구에서 각각 1개씩의 인구총조사구를 실사표본조사구로 추출하게 된다.

이해를 돕기 위하여 158개의 실사표본조사구를 추출하는 과정을 예로 들어 설명하기로 한다.

1) 實查標本調査區 抽出

우선 158개 조사구를 600개 기본표본조사구의 층별 분포에 비례하도록 각 층에 배분한다. 배분방법 및 결과는 <表 V-1>과 같다.

<表 V-1> 158個 調査區의 層別 配分結果

층	기본표본조사구수	계산절차 및 결과
(1) 6대 시의 동	231 (38.50)	$158 \times 38.50 / 100 = 60.83 = 61$
(2) 기타 시의 동	155 (25.83)	$158 \times 25.83 / 100 = 40.81 = 41$
(3) 시의 읍·면	86 (14.33)	$158 \times 14.33 / 100 = 22.64 = 23$
(4) 군의 읍	46 (7.67)	$158 \times 7.67 / 100 = 12.12 = 12$
(5) 군의 면	82 (13.67)	$158 \times 13.67 / 100 = 21.60 = 21$
합계	600(100.00)	158

이 表에서 각 층의 계산결과를 반올림하면 층별 합계가 159가 되므로 반올림된 것 중에서 소수점 이하의 수가 가장 적은 제5층의 21.60을 반올림하지 않고 21로 하였다. 제2표본의 층별 기본표본조사구수와 배분된 결과의 층별 기본표본조사구수를 정리하면 <表 V-2>와 같다.

<表 V-2> 第2標本 및 158個 基本標本調査區의 層別 比較

표본	계	1층	2층	3층	4층	5층
제2표본	200	77	52	29	15	27
실사표본	158	61	41	23	12	21

다음에는 제2표본의 각 층별 기본표본조사구 중에서 배분된 수의 기본표본조사구를 계통추출한다(바로 전 조사에서 제2표본이 사용되었다면 제3표본을 사용한다). 따라서 제1층에서는 77개 기본표본조사구 중에서 61개의 기본표본조사구를 계통추출하고, 나머지 층의 기본표본조사구 중에서도 배분된 수의 기본표본조사구를 계통추출한다.

여기서는 제4층의 15개 기본표본조사구 중에서 12개의 기본표본조사구를 계통추출하는 경우를 예로 들기로 한다.

우선 제4층에서의 추출간격을 계산하면 $15/12 = 1.25$ 가 된다. 다음에는 1.25 이하의 난수 하나를 택한다. 여기서는 0.78이 택해졌다고 하자. 이 0.78을 첫 번째 표본지정번호로 하고 여기에 추출간격을 계속 더한다. 이 때, 누적결과가 제4층의 원래의 기본표본조사구수인 15를 초과하면 누적을 중지한다. 그리고 누적결과의 정수부분만을 택한다. 다음은 표본지정번호 산출작업표의 작성결과이다. 이 작업표에는 이해를 돕기 위하여 출발번호가 1.05인 경우도 예시하였다.

순서	표 본 지 정 번 호	
	(1) 출발번호 = 0.78인 경우	(2) 출발번호 = 1.05인 경우
01	0.78 = 0.78 = 0	1.05 = 1.05 = 1
02	0.78 + 1.25 = 2.03 = 2	1.05 + 1.25 = 2.30 = 2
03	2.03 + 1.25 = 3.28 = 3	2.30 + 1.25 = 3.55 = 3
04	3.28 + 1.25 = 4.53 = 4	3.55 + 1.25 = 4.80 = 4
05	4.53 + 1.25 = 5.78 = 5	4.80 + 1.25 = 6.05 = 6
06	5.78 + 1.25 = 7.03 = 7	6.05 + 1.25 = 7.30 = 7
07	7.03 + 1.25 = 8.28 = 8	7.30 + 1.25 = 8.55 = 8
08	8.28 + 1.25 = 9.53 = 9	8.55 + 1.25 = 9.80 = 9
09	9.53 + 1.25 = 10.78 = 10	9.80 + 1.25 = 11.05 = 11
10	10.78 + 1.25 = 12.03 = 12	11.05 + 1.25 = 12.30 = 12
11	12.03 + 1.25 = 13.28 = 13	12.30 + 1.25 = 13.55 = 13
12	13.28 + 1.25 = 14.53 = 14	13.55 + 1.25 = 14.80 = 14
13	14.53 + 1.25 = 15.78 = 15	14.80 + 1.25 = 16.05 = 16

출발번호가 0.78인 경우를 보면, 마지막 누적결과가 15를 초과하였으므로 더 이상 누적하지 않았다. 소수점 이하의 수를 무시한 정수부분에 해당하는 번호의 기본표본조사구를 택하면 된다. 그런데 첫 번째의 표본지정번호는 0이므로 이는 해당이 되지 않는다. 즉 2, 3, 4, 5, 7, 8, 9, 10, 12, 13, 14, 15번째의 기본표본조사구를 택하면 된다.

출발번호가 1.05인 경우에는 1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14번의 기본표본조사구를 택하면 된다(표본지정번호 16의 경우에는 16번의 기본표본조사구가 없으므로 해당되지 않는다).

그런데 이와는 달리 15개 기본표본조사구 중에서 3개의 기본표본조사구를 제외하는 편법도 있다. 예를 들면, 이 때의 추출간격은 $15/3=5$ 가 된다. 출발번호로 5이하의 수로 2가 택해 졌다고 하면, 제외되는 경우의 표본지정번호는 2, 7, 12가 되므로 15개의 기본표본조사구 중에서 2, 7, 12번의 기본표본조사구를 제외한 나머지 12개의 기본표본조사구를 택하면 된다. 이 때, 출발번호가 1로 정해졌다고 하면 제외되는 기본표본조사구는 1, 6, 11번의 기본표본조사구가 되므로 표본지정번호 산출작업표에서 출발번호가 0.78인 경우의 결과와 같게 된다.

이와 같은 방법으로 추출된 158개의 기본표본조사구에서는 각각 1개씩의 인구총조사 조사구를 크기의 측도에 비례하는 확률로 추출한다.

이해를 돕기 위하여 제4층의 15개 기본표본조사구 중에서 추출된 12개 기본표본조사구 중의 2번째 기본표본조사구의 경우에 6개 인구총조사 조사구 중에서 1개의 인구총조사구를 실사표본조사구로 추출하는 경우를 예로 든다.

다음의 작업표에서 보는 바와 같이 우선 인구총조사구의 가구수를 10으로 나누어 반올림한 결과를 크기의 측도로 부여하고, 이를 누적한 결과는 39가 된다.

조사구 번호	가구수	크기의 측도	누적결과	표본지정번호
24-03-1	72	7	7	-
24-03-2	71	7	14	-
24-03-3	59	6	20	-
24-03-4	62	6	26	-
24-03-5	55	6	32	29
24-03-6	65	7	39	-

난수표 이용하여 39 이하의 수를 택한다. 여기서 29가 택해졌다고 하면, 누적 결과에서 29가 포함되는 조사구는 '24-03-5'번 조사구이므로 이 조사구를 실사표본조사구로 택한다.

만일 어떤 인구총조사구가 이미 다른 조사의 실사표본조사구로 사용되었다면 이 조사구는 제외하고 나머지 조사구 중에서 추출한다. 앞의 예에서 3번째 조사구가 이미 다른 조사에 사용되었다면, 이 조사구는 제외하고 크기의 측도를 누적하여 5개의 조사구 중에서 1개의 조사구를 크기의 측도에 비례하는 확률로 임의추출한다. 예를 들면 다음과 같다. 여기서는 크기의 측도의 최종 누적결과가 33이므로 그 이하의 수로 25가 택해진 것이다.

조사구 번호	가구수	크기의 측도	누적결과	표본지정번호
24-03-1	72	7	7	-
24-03-2	71	7	14	-
24-03-3	59	비해당	-	-
24-03-4	62	6	20	-
24-03-5	55	6	26	25
24-03-6	65	7	33	-

2) 實查標本調査區의 抽出確率

제1표본, 제2표본, 제3표본에서 각각 층별로 일부의 기본표본조사구를 계통추출하고, 추출된 각 기본표본조사구에서 실사표본조사구로 1개씩의 인구총조사구를 크기의 측도에 비례하는 확률로 추출하는 경우에 전체 인구총조사구 중에서 각 실사표본조사구가 추출된 확률은 다음의 4개 확률을 곱하면 된다.

- ① P_1 = 각 층에서 기본표본조사구가 추출된 확률
- ② P_2 = 각 층의 기본표본조사구가 제1표본, 제2표본, 제3표본으로 추출된 확률
- ③ P_3 = 각 표본에서 실사표본조사구가 포함된 기본표본조사구의 추출확률
- ④ P_4 = 각 기본표본조사구에서 실사표본조사구로 추출된 인구총조사구의 추출확률

따라서 이를 공식으로 표시하면 공식 (5.2)와 같이 된다. 이 공식에서는 공식 (5.1)에서 정의한 기호 이외에 다음의 기호가 추가되었다.

β = 제g표본의 h층에서 β 번째로 추출된 기본표본조사구를 나타내는 첨자 ($\beta = 1, 2, 3, \dots, n'$)

n' = 실사표본조사구가 추출되는 기본표본조사구수 (각 기본표본조사구에서 1개씩의 인구총조사구를 실사표본조사구로 추출하는 경우에는 실사표본조사구수와 같게 됨)

공식 : 실사표본조사구 추출확률 (2)

$$\begin{aligned}
 {}_gP_{h\beta j} &= P_{\text{①}} \times P_{\text{②}} \times P_{\text{③}} \times P_{\text{④}} \\
 &= \frac{n_h S_{hi}}{S_h} \cdot \frac{{}_g n_h}{n_h} \cdot \frac{{}_g n'_h}{{}_g n_h} \cdot \frac{{}_g S_{h\beta j}}{{}_g S_{h\beta}} \quad : S_{hi} = {}_g S_{h\beta} \\
 &= \frac{{}_g n'_h}{S_h} {}_g S_{h\beta j} \quad \dots\dots (5.2)
 \end{aligned}$$

이 공식에서 사용된 기호를 알기 쉽게 다른 방법으로 설명하면 다음과 같다.

${}_gP_{h\beta j}$ = 제g표본의 h번째 층에서 β 번째 기본표본조사구내의 j번째 인구총조사구의 추출확률

n_h = h번째 층의 기본표본조사구수

${}_g n_h$ = 제g표본에서 h번째 층의 기본표본조사구수

${}_g n'_h$ = 제g표본에서 h번째 층의 ${}_g n_h$ 개 기본표본조사구 중에서 추출된 기본표본조사구수

S_h = h번째 층의 크기의 측도(表 V-1 참조)

S_{hi} = h번째 층의 i번째 기본표본조사구의 크기의 측도

${}_g S_{h\beta}$ = 제g표본의 h번째 층에서 ${}_g n'_h$ 개 기본표본조사구 중의 β 번째 기본표본조사구의 크기의 측도

${}_g S_{h\beta j}$ = 제g표본의 h번째 층에서 ${}_g n'_h$ 개 기본표본조사구 중의 β 번째 기본표본조사구 내의 j번째 인구총조사구의 크기의 측도

공식 (5.2)를 (5.1)과 비교하여 보면 그 모양이 같다는 것을 알 수 있으며, 확률계산 방법이 원칙적으로 같다는 것도 알 수 있다.

제2표본의 제4층의 15개 기본표본조사구 중에서 추출된 12개 기본표본조사구 중의 2번째 기본표본조사구에서 5번째의 인구총조사구가 실사표본조사구로 추출된 경우에 전체 인구총조사구 중에서 이 실사표본조사구가 추출된 확률을 계산해 보기로 한다. 이 실사표본조사구의 인구총조사 당시의 일반가구수는 55가구이다.

우선 공식 (5.2)의 각 항목과 관계되는 값을 정리한 결과는 다음과 같다.

$$\begin{aligned} {}_g n'_h &= {}_2 n'_4 &&= 12 \\ {}_g S_{h\beta j} &= {}_2 S_{4, 02, 5} &&= 6 \text{ (55가구를 10으로 나누어 반올림)} \\ S_h &= S_4 &&= 61,178 \text{ (表 V-1 참조)} \end{aligned}$$

여기서 ${}_2 S_{4, 02, 5}$ 의 우측 첨자 4, 02, 5는 $h=4$, $\beta=2$, $j=5$ 를 3045로 표시하면 구분이 되지 않으므로, 이를 구분한 것이다.

위의 값을 공식 (5.2)에 대입하여 계산하면 다음과 같이 된다.

$$\begin{aligned} {}_g P_{h\alpha j} &= \frac{{}_g n'_h}{S_h} {}_g S_{h\alpha j} \\ {}_2 P_{4, 02, 5} &= \frac{12}{61,178} 6 = \frac{1}{5,098.1667} 6 \\ &= \frac{1}{849.6944} \end{aligned}$$

위에서 4층의 경우에는 분모 5,098.1667은 상수가 되므로 다른 실사표본조사구의 크기의 측도가 7이라면 분자의 6대신에 7로 바꾸어 계산하면 된다.

VI. 新築 아파트에 대한 措置

실사표본조사구 내에 아파트가 신축된 경우에는 신축 아파트는 제외하고 조사를 한다. 그 대신 전국의 신축 아파트를 대상으로 표본조사구를 추출하여 조사에 포함시켜야 한다.

여기서는 1997년도 전국 출산력 및 가족보건실태조사의 경우를 예상하여 신축 아파트에서 추가되는 표본조사구수 산출방법과 표본조사구 추출절차 및 방법을 예를 들어 설명하였다. 다른 조사의 경우에도 여기에서 택한 방법을 그대로 따르면 될 것이다.

1. 新築 아파트에 대한 標本調査區 抽出의 必要性

실사표본조사구 내에서는 모든 가구를 다 조사하는 것이 원칙이다. 그런데, 실사단계에서 조사구를 방문하여 보면 1995년 인구주택총조사(1995. 11. 1) 이후에 조사구 경계선 내에 주택의 신축, 특히 아파트의 신축으로 가구가 많이 증가한 경우가 있게 되는데, 이 경우에 조사구 경계선 안의 모든 가구를 다 조사하려면 주어진 조사기간에 조사를 완료하는 것이 불가능하게 된다.

이와 같은 경우에 우선적으로 생각할 수 있는 것이 조사기간의 연장이다. 그러나 조사기간을 연장하려면 그에 따른 추가소요 예산을 검토하여야 하는 등의 내부 행정처리가 먼저 이루어져야 하며, 그 다음 조사구부터 계속하여 조사일정이 변경되므로 시·도 및 구·시·군 보건소를 통하여 해당 동·읍·면에 조사일정 변경을 통보하여야 하는 등 업무절차가 복잡하게 된다.

또한 실제적으로 가구수가 너무나 많이 증가한 경우에는 표본오차가 증가하는 것은 감수하더라도 이 조사구의 조사기간을 1주일 이상 연장하여야 하는 경

우도 있게 되므로, 이러한 경우를 대비하여 두는 것이 바람직하다.

가구 증가의 경우 조사기간 연장을 피하는 방법으로는 조사구 안의 일부의 가구를 표본으로 추출하여 조사토록 하고 추정단계에서 가중치를 적용하는 방법이 있지만, 현지에서 지도원이 연구진의 지시를 받아 이러한 조치를 취하는 것도 쉽지가 않으며, 또 이와 같이 한다고 하더라도 표본오차는 여전히 증가하게 된다.

따라서 이러한 문제점을 조금이라도 해결하기 위하여 실사표본조사구 내의 신축 아파트는 조사에서 제외시키는 대신에 1995년 11월 1일 이후의 전국의 신축 아파트에 대하여는 별도의 표본조사구를 추출하여 조사하는 방법이 가장 바람직 하다. 신축 아파트 현황에 관한 자료는 통계청에서 얻을 수 있다.

신축 아파트를 제외한 기타 주택의 신축에 대하여는 자료수집이 불가능한 실정이므로, 실사표본조사구 경계선 내의 단독주택, 연립주택 및 다세대주택의 가구에 대하여는 모두 조사하거나 현지 실정에 따라 적절한 대책을 별도로 강구할 수밖에 없다.

2. 新築 아파트 現況資料의 整理

신축 아파트의 표본추출을 위한 관련자료는 다음과 같이 2개의 자료철로 보관되어 있다.

첫 번째 철 : 통계청 신축 아파트 현황

두 번째 철 : 신축 아파트의 층 및 지역별 가구수

첫 번째 철(통계청 신축 아파트 현황)에는 통계청에서 정리한 신축 아파트 현황 1차분(1995. 11. 1.~1996. 4. 31.)과 2차분(1996. 5. 1.~1996. 10. 31.)이 철해져 있으며, 1997년부터 매년 2번씩 통계청 자료를 추가하여야 한다.

통계청 통계기획국 조사관리과(표본2계)에서는 통계청의 가구단위 통계조사의 표본을 보완하기 위하여 매년 4월말과 10월말 기준으로 준공검사가 완료된 아파트에 관한 자료를 각 시도에서 제출받아 신축 아파트 현황표를 작성하고 있다. 다음의 〈樣式 3〉은 1996년도 2차분 신축 아파트 현황표의 첫 번째 부분의 일부를 예로 든 것이다.

〈樣式 3〉 新築 아파트現況 ('96. 5. 1.~'96. 10. 31.)

지 역	가구수	지 역	가구수	지 역	가구수
전 국	158,625	포천군	239	연기군	158,176
洞 部	135,309	김포군	242	당진군	569
邑 部	14,202	面 部	1,373	전라북도	6,383
面 部	9,214	평택시	329	洞 部	5,980
서 울	10,941	남양주시	817	전주시	3,048
洞 部	10,941	과주시	227	군산시	1,623
부 산	29,174	강 원 도	4,712	익산시	276
洞 部	29,174	洞 部	4,187	정읍시	1,033
대 구	13,234	강릉시	795	面 部	403
洞 部	10,938	동해시	1,220	정읍시	306
邑 部	1,364	태백시	210	무주군	40
面 部	932	속초시	1,438	임실군	57
		삼척시	524	전라남도	6,584

두 번째 철(신축 아파트의 층 및 지역별 가구수)에는 첫 번째 철의 '통계청 신축 아파트 현황'의 내용을 신축아파트 표본지역 추출용으로 사용하가 위하여 5개 층으로 나누어 정리한 결과가 철해져 있으며, 〈表 VI-1〉은 제5층의 첫부분을 예로 든 것이다.

〈表 VI-1〉 新築 아파트의 層 및 地域別 家口數 - 第5層

지역 분류	분류번호	1차	2차	계	누적
(5층)	50-00-00	2,561	4,014	6,575	-
경기 양주군 면부	50-31-01	-	250	250	250
경기 포천군 면부	50-31-02	-	299	299	549
경기 화성군 면부	50-31-03	-	49	49	598
충북 청원군 면부	50-33-01	1,003	557	1,560	2,158
충북 음성군 면부	50-33-02	130	-	130	2,288

이 表의 분류번호에서 좌측의 첫 번째 2단위는 층번호이며, 그 내용은 <表 VI-2>와 같다.

<表 VI-2> 新築 아파트의 層番號

층번호	내역	층번호	내역	층번호	내역
10	6대 시의 동	31	6대 시의 읍	34	기타 시의 면
20	시타 시의 동	32	기타 시의 읍	40	군의 읍
30	시부의 읍·면	33	6대 시의 면	50	군의 면

2단위의 층번호에서 첫 번째 단위만이 실제적인 층번호이며, 우측의 1단위는 제3층에만 적용되는 것으로 3층 내에서의 추가 분류번호이다.

따라서 층번호 30은 층번호 31~34까지를 통합한 것임을 뜻한다. 제3층이 아닌 다른 층의 경우에 우측의 '0'은 단위수를 2단위로 통일시키기 위하여 추가한 것 뿐이다.

두 번째의 2단위는 시·도번호이며, 그 내용은 <表 VI-3>과 같다.

<表 VI-3> 新築 아파트의 市·道番號

시도번호	시도	시도번호	시도	시도번호	시도	시도번호	시도
11	서울	24	광주	33	충북	37	경북
21	부산	25	대전	34	충남	38	경남
22	대구	31	경기	35	전북	39	제주
23	인천	32	강원	36	전남	00	전국

제2층~제5층의 경우에는 각 층 내에서 다시 시도별로 시·군명칭의 '가나다' 순서로 정리하고 도별로 시·군에 01번부터 일련번호를 부여하였다. 따라서 분류번호가 '50-31-02'인 경우에는 5층의 경기도에서 '가나다' 순서로 2번째의 군임을 나타낸다.

1997년부터 매년 2번씩 통계청의 신축 아파트 현황을 수집할 때마다 '신축 아파트의 층 및 지역별 가구수'도 현재와 같은 방법으로 보완하여야 한다.

3차분의 자료가 수집되는 경우 <表 VI-1>에서 '2차' 다음에 '3차'를 추가하고 이에 따라 3차분까지의 총가구수와 누적가구수를 수정하면 된다. 그리고 제2층~제5층의 경우에 새로운 시·군이 있을 때에는 '가나다' 순서로 해당 위치에 추가하고 시·군 일련번호를 수정하면 된다.

3. 新築 아파트의 實查標本調査區 抽出

1996년말 현재 계획으로는 1997년도 전국 출산력 및 가족보건실태조사의 실사표본조사구는 제2표본(200개 기본표본조사구)의 각 기본표본조사구에서 1개씩의 인구총조사구를 실사표본조사구로 추출하고 신축 아파트에 대하여는 별도의 실사표본조사구를 추출기로 하였다. 따라서 신축 아파트에 대한 실사표본조사구 추출방법을 1997년도 전국 출산력 및 가족보건실태조사의 경우를 예로 들어 설명하기로 한다.

가. 實查標本調査區數의 算出

1995년 인구총조사 이후 1996년 11월말까지의 신축 아파트 가구수는 309,816 가구이며, 크기의 측도는 $309,816/10 = 30,982$ 가 된다.

1995년 인구총조사구의 총 가구수는 12,885,650가구이며, 이 때 200개의 실사표본조사구를 추출하므로 같은 비율로 신축 아파트에서도 실사표본조사구를 추출한다면 신축 아파트 조사구수는 다음과 같이 산출된다.

$$\text{신축 아파트 실사표본조사구수} = 200 \frac{309,816}{12,885,650} = 4.81 = 5\text{개 조사구}$$

'신축 아파트의 층 및 지역별 가구수' 자료에서 각 층별 가구수만을 발췌하고, 1995년 인구총조사구를 이용하여 기본표본조사구를 추출할 때와 같이 읍부와 면부에 관련되는 제3~제5층에는 2배의 가중치를 적용하여 5개의 조사구를 각 층에 배분한 결과는 <表 VI-4>와 같다.

〈表 VI-4〉 層別 新築 아파트 調査區數 算出 內譯

층	가구수	가중치	② × ③	④의 구성비	⑤ × 5 조사구수
①	②	③	④	⑤	⑥
(1) 6대 시의 동	127,142	1	127,142	0.3594	1.80 = 2
(2) 기타 시의 동	138,712	1	138,712	0.3921	1.96 = 2
(3) 시의 읍면	26,067	2	52,134	0.1473	0.74 = 1
(4) 군의 읍	11,320	2	22,640	0.0640	0.32 = 0
(5) 군의 면	6,575	2	13,150	0.0372	0.18 = 0
계	309,816	-	353,778	1.0000	5.00 = 5

이 表의 ⑥란에서 보는 바와 같이 제4층과 제5층의 경우에는 1개의 조사구도 배분되지 않는다. 여기서 반올림하지 않고 올림을 한다면 4층과 5층에도 각각 1개 조사구가 배분된다.

그러나 표본오차의 산출을 위하여는 1개 층에 적어도 2개의 조사구는 있어야 하는데, 표본오차의 산출을 위하여 제3층, 제4층, 제5층의 조사구수를 각각 2개씩으로 하면 신축 아파트에서 추출되는 조사구수는 10개가 된다.

따라서 제4층, 제5층을 제3층에 통합시키고 제3층에서 2개의 표본조사구를 추출키로 하였다(원래의 제3층은 시의 읍·면지역이었으나 제4층과 제5층이 통합되었으므로 통합후에는 전국의 읍·면지역이 되며, 동지역과 구별이 된다).

나. 層別 標本 아파트의 抽出

우선 제1층(6대 시의 동), 제2층(기타 시의 동), 제3층(읍·면)에서 각각 2개씩의 표본조사구를 계통추출하기 위하여 각 층에서 2개씩의 표본지정번호를 선정하였으며, 그 결과는 〈表 VI-5〉와 같다.

이 表에서 추출간격은 가구수를 조사구수로 나눈 것이다. 표본지정번호 (1)은 출발번호로서 난수표에서 추출간격 이하인 수를 택한 것이며 표본지정번호 (2)는 출발번호에 추출간격을 더한 것이다.

〈表 VI-5〉 層別 標本指定番號 選定 內譯

층	가구수	조사구수	추출 간격	표본지정번호	
				(1)	(2)
(1) 6대 시의 동	127,142	2	63,571	23,418	86,989
(2) 기타 시의 동	138,712	2	69,356	57,304	126,660
(3) 읍·면	43,962	2	21,981	8,273	30,254

각 층에서의 실사표본조사구는 일차적으로 〈表 VI-5〉의 표본지정번호에 해당하는 지역을 ‘신축 아파트의 층 및 지역별 가구수’ 자료를 이용하여 추출한다.

‘신축 아파트의 층 및 지역별 가구수’ 자료에서 제1층의 경우를 예로 들면 〈表 VI-6〉과 같다(이 表에서 표본지정번호는 설명을 위하여 추가한 것이다).

〈表 VI-6〉 新築 아파트의 層 및 地域別 家口數 - 第1層

시도 및 지역 분류	분류번호	가구수			누적 가구수	표본 지정번호
		1차	2차	계		
〈제1층〉	10-00-00	67,608	59,534	127,142	-	-
서울 ---- 동부	10-11-01	10,941	24,428	35,369	35,369	23,418
부산 ---- 동부	10-21-01	29,174	10,562	39,736	75,105	-
대구 ---- 동부	10-22-01	10,938	7,706	18,644	93,749	86,989
인천 ---- 동부	10-23-01	6,516	6,195	12,711	106,460	-
광주 ---- 동부	10-24-01	6,801	6,321	13,122	119,582	-
대전 ---- 동부	10-25-01	3,238	4,322	7,560	127,142	-

〈表 VI-5〉에서 제1층의 표본지정번호는 23,418과 86,989이므로 이 表의 누적 가구수에서 23,418번째의 가구가 포함되는 지역을 찾으려면 서울이 된다. 그런데 서울의 1차분 가구수는 10,941가구이므로 표본지정번호 23,418은 1차분의 가구수를 초과한다.

따라서 1차분에 해당되지 않고 2차분에 해당된다. 2차분에서는 23,418(표본지

정번호) - 10,941(1차분 가구수) = 12,477이므로 2차분 신축 아파트 명부에서 서울지역의 12,477번째의 가구가 속하는 아파트를 추출한다.

다음에 2번째 표본지정번호는 86,989이므로 이 表의 누적가구수에서 86,989번째의 가구가 포함되는 지역은 대구이다. 그리고 대구에서 86,989(2번째 표본 지정번호) - 75,105(대구의 전 지역인 부산까지의 누적가구수) = 11,884번째의 가구는 1차분의 10,938가구를 초과하므로 1차분에는 해당하지 않고 2차분에 해당되며, 2차분에서 11,884 - 10,938(1차분 가구수) = 946이므로 2차분의 신축 아파트 명부에서 대구지역의 946번째 가구가 속하는 아파트를 추출한다.

제2층의 경우에는 제1층에서의 방법을 그대로 적용하면 되며, 결과적으로 충북 청주시 동부의 2차분 신축 아파트 명부에서 557번째 가구가 속하는 아파트와 경남 울산시 동부의 1차분 신축 아파트 명부에서 1,455번째 가구가 속한 아파트를 추출하면 된다.

제1층과 제2층에서 적용한 방법은 같지만 제3층의 경우에는 절차가 좀 복잡하다. <表 VI-5>에서 제3층의 표본지정번호는 8,273과 30,254이다. 그런데 제3층은 '신축 아파트의 층 및 지역별 가구수' 자료의 제3층, 제4층, 제5층이 통합된 것이므로 통합전의 각 층의 가구수를 정리한 결과는 <表 VI-7>과 같다.

<表 VI-7> 標本指定番號 決定 內譯 - 第3層

통합전의 층	분류번호	가구수			누적	표본 지정번호
		1차	2차	계		
(3) 시의 읍면	34-00-00	12,520	13,547	26,067	26,067	8,273
(4) 군의 읍	40-00-00	8,235	3,085	11,320	37,387	30,254
(5) 군의 면	50-00-00	2,561	4,014	6,575	43,962	-

이 表에서 통합전의 층, 분류번호, 가구수는 '신축 아파트의 층 및 지역별 가구수' 자료의 내용이며, 표본지정번호는 설명을 위하여 추가한 것이다.

그리고 가구수를 누적한 이유는 표본지정번호에 따른 해당 층을 결정하기 위한 것이다. 첫 번째의 표본지정번호는 8,273이므로 8,273번째의 가구는 통합전의 제3층에 포함되므로 통합전의 제3층에서는 8,273번째의 가구가 속하는 지역을 추출

한다. 즉, 통합전 제3층에서 표본지역 추출을 위한 표본지정번호는 8,273이 된다.

두 번째의 표본지정번호는 30,254이므로 30,254번째의 가구는 통합전의 제4층에 포함되므로, 통합전의 제4층에서 $30,254 - 26,067$ (제3층의 가구수) = 4,187번째의 가구가 속하는 지역을 추출한다. 즉 통합전 제4층에서 표본지역 추출을 위한 표본지정번호는 4,187이 된다.

통합전의 제3층과 제4층에서의 표본지역 선정과정을 설명하기 위하여 정리한 결과는 <表 VI-8>과 같다(이 表에서 표본지정번호는 설명을 위하여 추가한 것이다).

<表 VI-8> 統合前 第3層 및 第4層에서의 標本地域 抽出 內譯

통합전의 층 및 지역분류	분류번호	가구수			누적 가구수	표본 지정번호
		1차	2차	계		
제3층 경남 마산시 읍부	32-38-02	-	1,276	1,276	8,154	-
경남 밀양시 읍부	32-38-03	99	129	228	8,382	8,237
제4층 강원 홍천군 읍부	40-32-04	525	-	525	4,039	-
충북 단양군 읍부	40-33-01	422	-	422	4,461	4,187

통합전의 제3층에서 표본지역 추출을 위한 표본지정번호는 8,273이므로 8,273번째의 가구가 포함되는 지역은 경남 밀양시의 읍부이며, 여기서는 $8,273 - 8,154$ (전 지역까지의 누적가구수) = 119이므로 119번째의 가구가 속한 아파트를 추출하면 된다.

그런데 1차분의 가구수는 99가구이므로 1차분에는 해당되지 않으며 2차분에서 $119 - 99 = 20$ 번째의 가구가 속한 아파트를 추출하면 된다.

통합전 제4층의 경우에는 표본지역 추출을 위한 표본지정번호가 4,187이므로 4,187번째의 가구가 포함되는 지역은 충북 단양군의 읍부이며, 여기서는 $4,187 - 4,039$ (전 지역까지의 누적가구수) = 48이므로 48번째의 가구가 속한 아파트를 추출하면 된다. 따라서 1차분의 가구수는 422가구이므로 1차분에서 48번째의 가구가 속한 아파트를 추출하면 된다.

지금까지는 각 층에서 신축 아파트를 추출하여야 할 지역과 이 지역에서 아파트를 추출하여야 할 명부(1차분 또는 2차분)를 결정하고, 해당 신축 아파트 명부에서 아파트를 추출하기 위한 표본지정번호를 결정하였다.

다음으로는 실제로 아파트를 추출하여야 하는데 신축 아파트명부는 통계청에서 신축 아파트 현황과 함께 보관하고 있으므로, 통계청의 신축 아파트 명부를 열람하거나 담당자에게 의뢰하여 표본 아파트를 추출하여야 한다.

그러기 위해서는 <表 VI-9>의 자료가 있어야 하며, 이 表는 지금까지의 최종 결과를 정리한 것이다.

<表 VI-9> 新築 아파트 標本抽出을 위한 地域別 標本指定番號

지역	차수	표본 지정번호	지역	차수	표본 지정번호
(1) 서울 - 동부	2차분	3,477	(4) 경남 울산시 동부	1차분	1,455
(2) 대구 - 동부	2차분	946	(5) 경남 밀양시 읍부	2차분	20
(3) 충북 청주시 동부	2차분	557	(6) 충북 단양군 읍부	1차분	48

표본으로 추출되는 신축 아파트에 대하여는 아파트명, 소재지, 동수, 규모(가구수) 등의 자료가 필요하다. <表 VI-10>은 <表 VI-9>의 내용에 따라 통계청의 신축 아파트 명부를 열람하여 표본 아파트를 추출한 내용을 정리한 것이다.

<表 VI-10> 標本으로 抽出된 新築 아파트 資料

소재지	아파트명	동수	가구수	준공검사
(1) 서울 서대문구 홍제3동 455	현대그린아파트	2	240	'96. 7. 22.
(2) 대구 수성구 사월동 464	보성아파트	8	1,498	'96. 9. 7.
(3) 충북 청주시 상당구 울랑동 660-71	삼성아파트	8	774	'96. 5. 26.
(4) 경남 울산시 동구 전하동 550-42	현대한마음아파트	1	109	'95. 11. 6.
(5) 경남 밀양시 하남읍 수산리 489-14	동촌임대아파트	1	99	'96. 5. 29.
(6) 충북 단양군 단양읍 상진리 50-3	두진아파트	2	432	'96. 8. 26.

다. 實查標本調査區의 抽出

표본으로 추출된 아파트에 대하여는 동수, 동별 층수 및 가구수, 계단식 또는 복도식 여부 등의 관계자료를 수집한다. 이러한 자료의 수집방법으로는 해당 동·읍·면사무소에 문의하는 방법, 해당 아파트 관리사무소에 문의하는 방법, 현지출장으로 확인하는 방법 등이 있다.

자료가 수집된 각 아파트에서는 가구수에 의한 확률비례추출법으로 1개 동을 추출한다(1개 동뿐인 경우에는 그 동을 표본으로 추출하게 된다). 그리고 표본으로 추출된 동이 계단식인 경우에는 첫 번째 계단부터 60가구가 포함되는 조사구를 설정하고, 복도식인 경우에는 1층부터 60가구가 포함되는 조사구를 설정하여 실사표본조사구로 한다. 그리고 크기의 측도는 6을 부여한다.

예를 들어 <表 VI-10>의 (1)번 현대그린아파트는 101동과 102동의 2개 동이며, 각각 220가구로 출입구가 4개이고 15층이라고 할 때, 101동이 가구수에 의한 확률비례추출법으로 102동이 추출되었다면 102동의 1번 계단부터 2번 계단(1호출부터 4호출)까지를 실사표본조사구로 하면 된다.

VII. 標本設計 事例

여기서는 ① 1997년도 출산력 및 가족보건실태조사, ② 1998년도 국민건강조사, ③ 1998년도 보건의식행태조사, ④ 2000년도 장애인실태조사를 예상하여 작성한 표본설계개요(안)을 사례로 들었다. 국민건강 및 보건의식행태조사의 경우에 보건의식행태 부문은 국민건강조사의 조사 대상자 중의 일부를 다시 표본추출하여 조사하며 보고서도 별도로 발간하므로, 여기서는 분리하여 다루었다.

따라서 여기서는 이들 4개 조사의 보고서에 수록할 표본설계개요(안)을 미리 작성한 것이 되므로, 모든 수치는 가상 수치를 사용하였음을 밝혀 둔다. 따라서 그때 가서는 사실상의 수치로 수정하여 주면 될 것이다.

1. 1997年度 出産力 및 家族保健實態調査 標本設計概要(案)

가. 企劃段階에서의 決定事項

1) 調査母集團

가) 調査對象地域

각 시·도의 섬지역을 제외한 전국을 조사대상지역으로 하였다. 표본조사구는 전국을 대상지역으로 하여 추출하는 것이 바람직하지만, 교통이 불편한 섬지역에서 표본조사구가 추출되면 실제적으로 현지조사가 곤란하게 되는 경우가 있기 때문에 각 시·도의 섬지역은 조사대상지역에서 제외시켰다.

제외되는 섬지역의 가구수는 전국 가구수의 0.6%미만이므로 조사결과의 추정치에는 별다른 영향을 미치지 않는다.

나) 調査對象家口 및 家口員

조사대상가구에는 보통가구, 1인가구, 5인 이하의 비혈연가구가 포함되며, 조사대상 가구에서 통상적으로 생활을 같이 하는 사람을 조사대상 가구원으로 하였다. 그리고 부인조사의 경우에는 가구원중 15~49세의 기혼부인을 조사대상으로 하였다.

여기서 보통가구라 함은 원칙적으로 혈연관계가 있는 2명 이상이 생활을 같이 하고 있는 경우로서 일반적으로 말하는 가구의 개념이다 (비혈연 가구원이 함께 있는 경우도 있다).

1인가구는 혼자서 자취를 하는 경우와 다방, 상점, 음식점, 미장원 등에서 종업원 1명이 숙식을 하는 경우를 말하며, 비혈연가구는 혈연관계가 없는 5명 이하의 사람들이 함께 모여 가구의 형태를 유지하면서 사는 경우(예: 친구와 함께 자취를 하는 경우 및 다방, 음식점, 기숙사가 아닌 공장 내의 방에서 전혀 혈연관계가 없는 5명 이하의 종업원이 함께 숙식을 하는 경우)를 말한다.

보통가구 내에 혈연관계의 가구원 외에 비혈연관계의 사람도 함께 생활하고 있는 경우에는 3명까지는 보통가구의 가구원으로 취급하고 4~5명인 경우에는 별도의 비혈연가구로 분리시키기로 하였다.

2) 標本規模

조사규모는 1994년 조사의 조사구수(151개 조사구)와 가구수(11,100가구)보다는 크게 되도록 하되, 주어진 예산과 조사관리능력을 종합적으로 고려하여 조사구수는 210가구를 초과하지 않고 가구수는 13,000가구를 초과하지 않도록 하였다.

표본조사구는 1997~2001년까지 5년간 각종 가구단위 통계조사에 다목적으로 사용하기 위하여, 1995년 인구총조사 조사구 중에서 각 시·도의 섬지역 조사구와 시설단위 조사구를 제외한 보통조사구를 이용하여 1996년도에 추출한 기본표본조사구 중에서 추출키로 하였다.

기본표본조사구의 추출에 사용된 '인구총조사의 보통조사구'를 편의상 다음부터는 '인구총조사구'란 용어로 바꾸어 사용하기로 한다.

기본표본조사구는 150개 기본표본조사구의 제1표본, 200개 기본표본조사구의

제2표본, 250개 기본표본조사구의 제3표본으로 편성되어 있다. 기본표본조사구는 각각 5~7개(평균 5.6개)의 인구총조사구가 통합된 것이다. 이러한 기본표본조사구의 평균 가구수는 인구총조사의 가구수로 350가구이며 기본표본조사구에 포함된 인구총조사구의 평균 가구수는 63가구이다.

실사표본조사구수는 210개 조사구를 초과할 수 없으므로 제2표본의 200개 기본표본조사구에서 각각 1개씩의 인구총조사구를 실사표본조사구로 추출하면 된다.

그리고 실사표본조사구로 추출된 인구총조사구 내에 1995년 인구총조사 이후에 주택이 많이 신축되었을 때(특히 아파트가 신축되었을 때)에는 가구수가 많이 증가되어도 경계선 안의 모든 가구를 전부 조사하여야 한다.

그러나 가구수가 너무 많이 증가한 경우에는 모든 가구를 전부 조사하는 것이 현실적으로 곤란하게 되는 경우가 있어, 자료수집이 가능한 1995년 인구총조사 이후의 신축 아파트에 대하여는 별도로 표본조사구를 추출하여 추가기로 하였다.

인구총조사구의 총 가구수는 12,885,650가구이며, 1995년 인구총조사(1996. 11. 1)이후 1996년 10월말까지의 신축 아파트 자료를 수집하여 정리한 결과의 전체 규모는 309,816가구분이므로 가구수의 비는 100.0:2.4가 된다. 따라서 이들 두 자료에서 같은 비율로 실사표본조사구를 추출한다면 신축 아파트에서는 5개($2.4 \times 2 = 4.8$)의 조사구를 추출하여야 한다.

결과적으로 실사표본조사구는 기본표본조사구의 제2표본의 200개 기본표본조사구에서 각각 1개씩의 인구총조사구를 추출하고 신축 아파트에서는 5개의 표본조사구를 추출기로 하였다(실제적으로는 신축 아파트의 표본조사구 추출과정에서 6개 조사구로 조정됨).

나. 基本標本調査區 抽出概要

1) 基礎資料의 整理

1995년 인구총조사의 잠정집계에서 사용된 조사구별 자료에서 각 시·도의 섬지역 조사구와 시설단위 조사구를 제외한 219,771개 인구총조사구 자료로 전산작업화일을 작성하였다. 제외된 섬지역의 가구수는 실제적으로 전국 가구수의

0.6% 미만에 불과하다.

2) 人口總調查區의 層化

인구총조사구를 행정구역에 따라 다음과 같이 5개 층으로 층화하였다.

제1층 : 6대 시(서울, 부산, 대구, 인천, 광주, 대전)의 동

제2층 : 기타 시의 동

제3층 : 시의 읍·면

제4층 : 군의 읍

제5층 : 군의 면

이 때, 1995년 인구총조사 이후 1996년 7월말까지 행정구역이 개편되어 면에서 읍으로 승격되거나 군에서 시로 편입된 읍·면의 인구총조사구는 신행정구역에 따라 층번호를 수정하였다.

3) 層別 人口總調查區의 整列 및 크기의 測度 賦與

제3층(시의 읍·면)을 제외한 제1층, 제2층, 제4층, 제5층의 4개 층에서는 각각 인구총조사구를 그 주택특성에 따라 (1) 단독주택이 많은 조사구, (2) 아파트가 많은 조사구, (3) 연립 및 다세대주택이 많은 조사구, (4) 기타 조사구의 순서로 분류하였으며, 각 분류 내에서는 행정구역번호 및 조사구번호 순서로 정렬하였다.

제3층(시의 읍·면)의 경우에는 우선 (1) 6대 시의 읍, (2) 기타 시의 읍, (3) 6대 시의 면, (4) 기타 시의 면과 같은 순서가 되도록 먼저 분류한 다음, 각 분류 내에서는 앞서서와 같이 주택특성으로 재분류하고, 다시 행정구역번호 및 조사구번호 순서로 정렬하였다.

각 층별로 인구총조사구를 이와 같이 정렬한 다음에는 인구총조사구별로 일반가구수를 10으로 나누어 반올림한 결과를 크기의 측도로 부여하였다.

결과적으로 219,771개 인구총조사구의 전체 가구수는 12,885,650가구이었으며, 크기의 측도 부여결과의 전체 크기의 측도는 1,299,167이 되었다. 이에 관한 층별 내역은 <表 VII-1>과 같다.

이 表에서 각 층의 가구수를 10으로 나눈 결과와 크기의 측도가 비슷하지 않은 것은 가구수의 끝단위가 5~9인 조사구가 우연적으로 많기 때문이다.

〈表 VII-1〉 層別 크기의 測度

층	인구총조사구수	가구수	크기의 측도
(1) 6대 시의 동	104,222 (47.4)	6,035,168 (46.8)	608,395 (46.83)
(2) 기타 시의 동	68,589 (31.2)	4,051,452 (31.5)	408,580 (31.45)
(3) 시의 읍·면	18,825 (8.6)	1,120,502 (8.7)	112,966 (8.69)
(4) 군의 읍	10,220 (4.6)	606,878 (4.7)	61,178 (4.71)
(5) 군의 면	17,915 (8.2)	1,071,650 (8.3)	108,048 (8.32)
합계	219,771(100.0)	12,885,650(100.0)	1,299,167(100.00)

4) 層別 基本標本調査區數의 決定

기본표본조사구의 수를 600개 조사구로 하고, 층별 크기의 측도에 비례하도록 배분한 결과는 〈表 VII-2〉의 제1안과 같다.

〈表 VII-2〉 層別 基本標本調査區數 配分 內譯

층	인구총조사		기본표본조사구수	
	조사구수	크기의 측도	제1안	제2안
(1) 6대 시의 동	104,222	608,395 (46.83)	281 (46.8)	231 (38.5)
(2) 기타 시의 동	68,589	408,580 (31.45)	189 (31.5)	155 (25.8)
(3) 시의 읍·면	18,825	112,966 (8.69)	52 (8.7)	86 (14.3)
(4) 군의 읍	10,220	61,178 (4.71)	28 (4.7)	46 (7.7)
(5) 군의 면	17,915	108,048 (8.32)	50 (8.3)	82 (13.7)
계	219,771	1,299,167(100.00)	600(100.0)	600(100.0)

현실적으로 대부분의 조사에서는 동부와 읍·면부의 별도 추정치가 요구될 것이지만, 100개 조사구를 조사하는 경우에 읍·면부(3층+4층+5층)의 조사구수는 22개 조사구밖에 되지 않으며, 150개 조사구를 조사하는 경우에는 33개 조사구, 200개 조사구를 조사하는 경우에는 43개 조사구밖에 되지 않아 기본적으로 조

사구수가 적기때문에 읍·면부 추정치의 표본오차가 크게 되어, 분석단계에서 읍·면부에 대하여 언급하는 것은 별다른 의미를 찾을 수 없게 된다.

이와 같은 문제점을 조금이나마 해소하기 위하여 동지역(제1층과 제2층)의 크기의 측도는 그대로 두고 읍·면지역(제3~제5층)의 크기의 측도는 2배를 한 결과에 비례배분하였으며, 배분 결과는 <表 VII-2>의 제2안과 같다.

제2안에서 100개 조사구를 조사하는 경우에 읍·면부의 조사구수는 36개 조사구가 되며 150개 조사구를 조사하는 경우에는 54개 조사구, 200개 조사구를 조사하는 경우에는 71개 조사구가 되어 제1안의 경우보다는 제2안의 경우가 합리적이다. 따라서 층별 기본표본조사구의 수는 제2안을 택하였다.

5) 基本標本調査區의 抽出

<表 VII-2>의 제2안의 내용대로 6대 시의 동에서 231개, 기타 시의 동에서 155개, 시의 읍·면에서 86개, 군의 읍에서 46개, 군의 면에서 82개의 인구총조사구를 크기의 측도에 의한 확률비례계통추출법으로 추출하였다.

그리고 표본으로 추출된 인구총조사구부터 크기의 측도의 합계가 32를 초과할 때까지 그 다음의 인구총조사구를 통합하여 기본표본조사구로 하였다. 따라서 기본표본조사구의 인구총조사 가구수는 320가구 이상이 되었다.

이와 같이 기본표본조사구를 추출한 결과를 요약하면 <表 VII-3>과 같다.

<表 VII-3> 基本標本調査區 抽出結果

층	기본표본 조사구수	인구총조사구수	크기의 측도	가구수
(1) 6대 시의 동	231 (38.50)	1,323 (36.37)	4,738 (38.09)	81,168 (38.21)
(2) 기타 시의 동	155 (25.83)	863 (20.63)	2,729 (21.94)	54,602 (21.95)
(3) 시의 읍·면	86 (14.33)	479 (36.37)	4,738 (38.09)	30,160 (38.21)
(4) 군의 읍	46 (7.67)	239 (20.63)	2,729 (21.94)	16,185 (21.95)
(5) 군의 면	82 (13.67)	459 (43.00)	4,971 (39.97)	28,693 (39.84)
계	600(100.00)	3,363(100.00)	12,438(100.00)	210,808(100.00)

6) 基本標本調査區의 分類

600개의 기본표본조사구를 150개 조사구(제1표본), 200개 조사구(제2표본), 250개 조사구(제3표본)의 3종류로 분류하기 위하여, 600개 전체 기본표본조사구의 층별 구성비에 비례하도록 150개 조사구, 200개 조사구, 250개 조사구를 각각 층에 배분하였으며, 그 결과는 <表 VII-4>와 같다.

<表 VII-4> 標本の 種類別 基本標本調査區의 層別 配分內譯

층	기본표본조사구	제1표본	제2표본	제3표본
(1) 6대 시의 동	231 (38.50)	58 (38.67)	77 (38.50)	96 (38.40)
(2) 기타 시의 동	155 (25.83)	38 (25.33)	52 (26.00)	65 (26.00)
(3) 시의 읍·면	86 (14.33)	21 (14.00)	29 (14.50)	36 (14.40)
(4) 군의 읍	46 (7.67)	12 (8.00)	15 (7.50)	19 (7.60)
(5) 군의 면	82 (13.67)	21 (14.00)	27 (13.50)	34 (13.60)
계	600(100.00)	150(100.00)	200(100.00)	250(100.00)

그리고 각 층별로 전체 기본표본조사구 중에서 제1표본의 기본표본조사구를 첫 번째로 계통추출하였으며, 제2표본의 기본표본조사구를 두 번째로 계통추출하였다. 따라서 나머지의 기본표본조사구를 제3표본으로 하였다.

다. 實查標本調査區의 抽出

1) 第2標本에서의 實查標本調査區 抽出

제2표본의 200개 기본표본조사구에서 각각 1개씩의 인구총조사구를 크기의 층도에 따른 확률비례추출법으로 추출하여 실사표본조사구로 하였으며, 200개 실사표본조사구의 인구총조사 가구수는 12,109가구가 되었다.

2) 新築 아파트에서의 實查標本調査區 抽出

1995년 인구총조사(1995. 11. 1)이후 1996년 10월말까지의 신축 아파트 자료를

수집하여 집계한 결과의 전체 규모는 309,816가구분이었다. 이 신축 아파트를 6대 시의 동(제1층), (2) 기타 시의 동(제2층), (3) 읍·면(제3층)의 3개 층으로 층화하였다. 그리고 인구총조사구에서 기본표본조사구를 추출할 때와 같이 동(제1층과 제2층)의 가구수는 그대로 두고 읍·면(제3층)의 가구수에는 2배를 한 결과에 5개의 조사구를 비례배분하였으며, 그 결과는 <表 VII-5>와 같다.

<表 VII-5> 層別 新築 아파트의 實查標本調査區數 決定 內譯

층	가구수	크기의 측도	조사구수
(1) 6대 시의 동	127,142 (41.04)	12,714	1.8 → 2
(2) 기타 시의 동	138,712 (44.77)	13,871	2.0 → 2
(3) 읍·면	43,962 (14.19)	4,396	1.2 → 2
계	309,816(100.00)	30,981	5.0 → 6

이 表에서 보는 바와 같이 제3층의 경우에는 반올림을 하면 1개 조사구밖에 되지 않는다. 그런데 표본오차의 산출을 위하여는 1개 층의 표본조사구수는 적어도 2개 조사구 이상이 되어야 하므로, 제3층의 경우에는 1개 조사구를 2개 조사구로 조정하여 총 6개 조사구로 하였다.

제1층(6대 시의 동)에서는 신축 아파트를 서울, 부산, 대구, 인천, 광주, 대전의 순서로 분류하였으며, 제2층(기타 시의 동)의 경우에는 경기, 강원, ..., 제주의 순서로 분류하였다. 그리고 읍·면(제3층)에서는 6대 시의 읍, 기타 시의 읍, 6대 시의 면, 기타 시의 면, 군의 읍, 군의 면 순서로 먼저 분류하고 각 분류 내에서는 서울, 부산, ..., 대전, 경기, ..., 제주의 순서로 재분류하였다.

각 층에서는 크기의 측도에 의한 확률비례계통추출법으로 2개씩의 아파트를 추출하였다. 표본으로 추출된 각 아파트에서는 크기의 측도에 의한 확률비례추출법으로 1개의 동을 추출하였다. 그리고 표본으로 추출된 동이 계단식인 경우에는 첫 번째 계단부터, 복도식인 경우에는 1층부터 시작하여 60호가 포함되는 조사구를 설정하여 실사표본조사구로 하였다. 이와 같이 추출한 결과의 6개 아파트 조사구의 가구수는 360가구가 되었다.

3) 實查標本調査區 抽出結果의 標本規模

이와 같이 206개 실사표본조사구를 추출한 결과의 총가구수는 12,469가구가 되었으며, 그 세부 내역은 <表 VII-6>과 같다.

<表 VII-6> 206個 實查標本調査區 抽出結果의 層別 內譯

자료 및 층	모집단의 크기의 측도	실사표본조사구			표본비율
		조사구수	크기의 측도	가구수	
인구총조사구					
(1) 6대 시의 동	608,395	77	467	4,670	1/1,303
(2) 기타 시의 동	408,580	52	315	3,133	1/1,297
(3) 시의 읍·면	112,966	29	174	1,754	1/ 649
(4) 군의 읍	61,178	15	90	905	1/ 680
(5) 군의 면	108,048	27	166	1,647	1/ 651
소계	1,299,167	200	1,212	12,109	1/1,072
신축 아파트					
(6) 6대 시의 동	12,714	2	12	120	1/1,060
(7) 기타 시의 동	13,871	2	12	120	1/1,156
(8) 읍·면	4,396	2	12	120	1/ 366
소계	30,981	6	36	360	1/ 861
동	1,043,560	133	806	8,043	1/1,156
읍·면	286,588	73	442	4,426	1/ 366
계	1,330,148	206	1,248	12,469	1/1,066

여기서 인구총조사 이후의 가구증가율을 고려하면 조사때에는 약 12,700가구가 될 것으로 보았다. 그리고 조사불응, 장기 출타 및 기타 사유에 의한 조사미완률을 5.3%로 보면 약 12,000가구에 대하여 가구기본조사를 완료하게 되며, 가구기본조사 완료가구 내의 조사대상 기혼부인수는 약 7,400명이 될 것으로 예상된다. 부인조사의 조사미완률을 7%로 한다면 약 6,900명에 대하여 조사를 완료할 수 있을 것으로 기대하였다.

라. 調查結果의 家口數 및 婦人數

206개 표본조사구에서 전수조사한 결과의 총가구수는 12,670가구였으며, 이 중에서 12,000가구(94.7%)에 대하여 가구조사를 완료하였다.

그리고 가구조사가 완료된 가구의 총 대상기혼부인수는 7,400명이었으며, 이 중에서 6,810명(92.0%)에 대하여 부인조사를 완료하였다. 이에 관한 내역은 <表 VII-7>과 같다.

<表 VII-7> 調查結果 家口數 및 婦人數

지역	조사구수	가구수	조사완료 가구수	완료가구내 부인수	조사완료 부인수
동	133	8,270	7,690 (93.0)	5,390	4,900 (91.0)
읍·면	73	4,400	4,310 (98.0)	2,010	1,910 (95.0)
계	206	12,670	12,000 (94.7)	7,400	6,810 (92.0)

마. 推定方法

이 조사결과에 의한 각종 통계치는 가중표본합계치(weighted sample total)와 이를 기초로 한 평균 및 구성비이다.

여기서 단순표본합계치(simple sample total)를 사용하지 않고 가중표본합계치를 사용한 이유는 각 표본조사구가 표본으로 추출된 확률이 동일하지 않고 조사완료율이 표본조사구마다 다르기 때문에 이를 반영하기 위한 것이다.

가구조사 결과에서 어떤 특성 x 를 갖는 인구수의 가중표본합계치는 조사완료된 가구수를 기준으로 다음과 같이 산출한다.

$$X' = \sum_h \sum_i M'_{hi} X_{hi}$$

$$M'_{hi} = M_{hi} \frac{A}{\sum_h \sum_i M_{hi} A_{hi}}$$

$$M_{hi} = \frac{S_h}{n_h S_{hi}} \frac{A'_{hi}}{A_{hi}}$$

위의 공식에서

X' = 특성 x 를 갖는 인구수의 가중표본합계치

X = 특성 x 를 갖는 것으로 조사된 인구수

M' = 가중표본합계치 산출용 승수

M = 조사모집단의 추정 가구수 산출용 승수

A' = 가구수(조사완료 + 조사미완)

A = 조사완료 가구수

n = 표본조사구수

S = 크기의 측도

h = 층을 나타내는 첨자

i = 표본조사구를 나타내는 첨자

그리고 기혼부인조사 결과에서 어떤 특성 y 를 갖는 부인수 및 관련 변수의 가중표본합계치는 조사완료된 기혼부인수를 기준으로 다음과 같이 산출한다.

$$Y' = \sum_h \sum_i e M'_{hi} Y_{hi}$$

$$e M'_{hi} = e M_{hi} \frac{E}{\sum_h \sum_i e M_{hi} E_{hi}}$$

$$e M_{hi} = M_{hi} \frac{E'_{hi}}{E_{hi}}$$

위의 공식에서,

Y' = 특성 y 를 갖는 기혼부인수의 가중표본합계치

Y = 특성 y 를 갖는 것으로 조사된 기혼부인수

E' = 기혼부인수(조사완료+조사미완)

E = 조사완료 기혼부인수

e = 기혼부인에 관한 승수임을 나타내는 첨자

2. 1998年度 國民健康調査 標本設計概要(案)

가. 企劃段階에서의 決定事項

1) 調査母集團

각 시·도의 섬지역을 제외한 전국을 조사대상지역으로 하였다. 표본조사구는 전국을 대상지역으로 하여 추출하는 것이 바람직하지만, 교통이 불편한 섬지역에서 표본조사구가 추출되면 현실적으로 조사가 곤란하게 되는 경우가 있으므로 각 시·도의 섬지역은 조사대상지역에서 제외시켰다(제외된 섬지역의 총 가구수는 전국 가구수의 0.6% 미만임).

조사대상가구에는 보통가구와 1인가구를 포함시켰다. 여기서 보통가구라 함은 원칙적으로 혈연관계가 있는 2인 이상이 사람들이 생활을 같이 하고 있는 경우로서 일반적으로 말하는 가구의 개념이며, 비혈연 가구원이 함께 있는 경우도 있다. 그리고 1인가구는 혼자서 자취를 하는 경우와 다방, 상점, 음식점, 미장원 등에서 종업원 1명이 혼자서 숙식을 하는 경우 등을 말한다.

일반적인 가구조사의 경우에는 친구와 함께 자취를 하는 경우와 다방, 음식점, 공장 내의 방에서 혈연관계가 없는 종업원이 함께 숙식을 하는 경우 등의 비혈연가구도 포함되나, 본 조사에서는 비혈연가구의 경우에 가구원 서로가 다른 사람의 의료이용, 이환 등에 관한 사항을 알지 못하는 경우가 대부분이기 때문에 현실적으로 조사가 불가능하므로 조사대상 가구에서 제외시켰다.

이와 같은 이유로 보통가구의 가구원 중에서도 비혈연 가구원은 조사대상 가구원에서 제외시켰다.

2) 標本規模

조사규모는 1995년도 조사의 조사구수(110개 조사구)와 가구수(7,700가구)보다는 크게 되도록 하되, 주어진 예산과 조사관리능력 등을 고려하여 조사구수는 160개 조사구를, 그리고 가구수는 10,000가구를 초과하지 않도록 하였다.

표본조사구는 1997~2001년까지 5년간 각종 가구단위 통계조사에 다목적으로

사용하기 위하여, 1995년 인구총조사 조사구 중에서 각 시·도의 섬지역 조사구와 시설단위 조사구를 제외한 보통조사구를 이용하여 1996년도에 추출한 기본표본조사구 중에서 추출기로 하였다.

기본표본조사구의 추출에 사용된 ‘인구총조사의 보통조사구’를 편의상 다음부터는 ‘인구총조사구’란 용어로 바꾸어 사용하기로 한다.

기본표본조사구는 150개 기본표본조사구의 제1표본, 200개 기본표본조사구의 제2표본, 250개 기본표본조사구의 제3표본으로 편성되어 있으며, 각 기본표본조사구는 5~7개(평균 5.6개)의 인구총조사구가 통합된 것이다. 이러한 기본표본조사구의 평균 가구수는 인구총조사의 가구수로 350가구이며, 기본표본조사구에 포함된 인구총조사구의 평균 가구수는 63가구이다.

실사표본조사구는 160개 조사구를 초과할 수 없으므로 제1표본의 150개 기본표본조사구에서 각각 1개씩의 인구총조사구를 추출하여 실사표본조사구로 하된다. 그런데 표본으로 추출된 인구총조사구 내에 1995년 인구총조사 이후에 주택이 많이 신축되었을 때(특히 아파트가 신축되었을 때)에는 가구수가 많이 증가되어도 경계선 안의 모든 가구를 전부 조사하여야 하지만, 가구수가 너무 많이 증가한 경우에는 모든 가구를 전부 조사하는 것이 현실적으로 곤란하게 되는 경우가 있어, 자료수집이 가능한 1995년 인구총조사 이후의 신축 아파트에 대하여는 별도로 표본조사구를 추출하여 추가기로 하였다.

인구총조사구의 총 가구수는 12,885,650가구이며, 1995년 인구총조사(1995. 11. 1) 이후 1997년 10월말까지의 신축 아파트 가구수는 약 62,000가구분으로 예상되므로 가구수의 비는 100.0 : 4.8이 된다. 따라서 이들 두 자료에서 같은 비율로 실사표본조사구를 추출한다면 인구총조사구에서 150개의 조사구를 추출하게 되므로 신축 아파트에서는 7개($4.8 \times 150 / 100 = 7.2$)조사구를 추출하여야 한다.

결과적으로 실사표본조사구는 기본표본조사구의 제1표본의 150개 기본표본조사구에서 각각 1개씩의 인구총조사구를 실사표본조사구로 추출하고 신축 아파트에서는 60가구 기준의 7개 실사표본조사구를 추출기로 하였다(실제적으로는 신축 아파트의 조사구 추출과정에서 8개 조사구로 조정됨).

이와 같이 157개의 실사표본조사구를 추출하면 자료상의 조사구 내의 가구수(150개 인구총조사구의 가구수 + 7개 신축 아파트조사구의 가구수)는 약 9,600

가구가 될 것이며, 1995년 인구총조사 이후의 가구증가를 고려하면 실제 조사 때에는 9,700가구가 될 것으로 보았다. 그리고 불응, 장기 출타 및 기타 사유에 의한 조사미완률을 4%로 보면 9,310가구에 대하여 가구기본조사를 완료하게 되며, 1995년도 조사결과와 대상외의 가구율(비혈연가구율) 0.8%를 고려하면 대상 가구는 9,240가구가 될 것으로 기대하였다.

나. 基本標本調査區 抽出概要

이 부분은 '1997년도 출산력 및 가족보건실태조사 표본설계개요(안)'에서 설명한 '기본표본조사구 추출개요'와 같으므로 여기서는 생략한다.

다. 實查標本調査區의 抽出

1) 第2標本에서의 實查標本調査區 抽出

제1표본의 150개 기본표본조사구에서 각각 1개씩의 인구총조사구를 크기의 측도에 따른 확률비례추출법으로 추출하여 실사표본조사구로 하였다.

결과적으로 150개 실사표본조사구의 1995년 인구총조사 가구수는 9,084가구가 되었으며, 크기의 측도는 909가 되었다.

2) 新築 아파트에서의 實查標本調査區 抽出

인구총조사구의 총 가구수는 12,885,650가구이며, 1995년 인구총조사(1995. 11. 1) 이후 1997년 10월말까지의 신축 아파트의 전체 규모는 619,632가구분이므로 가구수의 비는 100.0 : 4.8이 된다. 따라서 이들 두 자료에서 같은 비율로 실사표본조사구를 추출한다면 인구총조사구에서는 150개의 조사구를 추출하므로 신축 아파트에서는 7개($4.8 \times 150 / 100 = 7.2$) 조사구를 추출하여야 한다.

619,632가구분의 신축아파트를 6대 시의 동(제1층), (2) 기타 시의 동(제2층), (3) 읍·면(제3층)의 3개 층으로 층화하였다.

그리고 인구총조사구에서 기본표본조사구를 추출할 때와 같이 동(제1층 및 제2층)의 가구수는 그대로 두고 읍·면(제3층)의 가구수에는 2배를 하여 7.2개의

조사구를 각 층에 배분하고, 배분결과를 반올림하여 8개 조사구로 조정하였다. 이에 관한 세부 내역은 <表 VII-8>과 같다.

<表 VII-8> 層別 新築 아파트의 實查標本調査區數 決定 內譯

층	가구수	크기의 측도	조사구수
(1) 6대 시의 동	254,284	25,429 (41.04)	2.6 → 3
(2) 기타 시의 동	277,424	27,742 (44.77)	2.8 → 3
(3) 읍·면	87,924	8,792 (14.19)	1.8 → 2
계	619,632	61,963(100.00)	7.2 → 8

표본조사구의 추출을 위하여 제1층(6대 시의 동)에서는 신축 아파트를 행정구역번호에 따라 서울, 부산, 대구, 인천, 광주, 대전의 순서로 분류하였으며, 제2층(기타 시의 동)의 경우에는 경기, 강원, ..., 제주의 순서로 분류하였다.

제3층(읍·면)의 경우에는 6대 시의 읍, 기타 시의 읍, 6대 시의 면, 기타 시의 면, 군의 읍, 군의 면과 같은 순서가 되도록 먼저 분류하고 각 분류 내에서는 서울, 부산, ..., 대전, 경기, ..., 제주의 순서로 재분류하였다.

그리고 제1층과(6대 시의 동)과 제2층(기타 시의 동)에서는 3개씩의 아파트를, 제3층(읍·면)에서는 2개의 아파트를 가구수에 의한 확률비례계통추출법으로 추출하였으며, 표본으로 추출된 각 아파트에서는 크기의 측도에 의한 확률비례추출법으로 1개의 동을 추출하였다.

표본으로 추출된 동이 계단식인 경우에는 첫 번째 계단부터 60가구가 포함되는 조사구를 설정하여 실사표본조사구로 하였으며, 복도식인 경우에는 1층부터 시작하여 60가구가 포함되는 조사구를 설정하여 실사표본조사구로 하였다. 이때의 조사구별 크기의 측도는 6으로 하였다. 결과적으로 신축 아파트에서 추출된 8개 조사구의 가구수는 480가구가 되었으며, 크기의 측도는 48이 되었다.

3) 實查標本調査區 抽出結果의 標本規模

158개 실사표본조사구를 추출하여 가구수를 집계한 결과, 총 가구수는 9,564

가구가 되었으며, 크기의 측도는 957이 되었다. 이에 관한 세부 내역은 <表 VII-9>와 같다.

<表 VII-9> 實查標本調査區 抽出結果의 標本規模

자료 및 층	모집단의 크기의 측도	실사표본			표본비율
		조사구수	크기의 측도	가구수	
인구총조사구					
(1) 6대 시의 동	608,395	58	350	3,503	1/1,738
(2) 기타 시의 동	408,580	38	235	2,350	1/1,739
(3) 시의 읍·면	112,966	21	132	1,316	1/ 856
(4) 군의 읍	61,178	12	68	679	1/ 900
(5) 군의 면	108,048	21	124	1,236	1/ 871
소계	1,299,167	150	909	9,084	1/1,429
신축 아파트					
(6) 6대 시의 동	25,429	3	18	180	1/1,413
(7) 기타 시의 동	27,742	3	18	180	1/1,541
(8) 읍·면	8,792	2	12	120	1/ 733
소계	61,963	8	48	480	1/1,291
동	1,070,146	102	621	6,213	1/1,723
읍·면	290,984	56	336	3,351	1/ 866
계	1,361,130	158	957	9,564	1/1,422

라. 調査結果 家口數 및 人口數

각 조사구에서 3명의 조사원은 지도원이 정하여 준 담당 구역 내의 가구를 거쳐 단위로 순서대로 방문하면서 가구명부에 각 가구에 관한 사항을 기입하고 가구 기본조사표를 먼저 작성하였다. 그리고 가구기본조사표의 내용에 따라 해당 개인조사표(이환조사표, 외래이용조사표, 입원이용조사표)를 작성하였다.

그런데, 주택의 신축으로 가구가 많이 증가한 조사구, 낮에는 응답자를 만날 수 없는 경우가 특히 많은 조사구, 불응가구가 많아 설득에 많은 시간이 소요되

는 조사구에서는 모든 가구를 다 조사하려면 조사기간을 연장하여야 한다.

그러나 조사기간 연장에 수반되는 추가예산이 고려되어야 하며, 예산이 가능하더라도 다음 조사구가 속한 모든 동·읍·면에 조사기간 변경사실을 통보하여야 하는 등의 행정절차가 복잡하게 된다. 따라서 각 조사구별로 정해진 기간 내에 조사를 위하여 방문하지 못하게 된 가구는 연구책임자의 허락을 받아 가구명부에만 기록토록 하였다.

현지조사 결과, 158개 표본조사구의 가구명부에 기입된 총가구수는 9,741가구가 되었으며, 이 중에서 가구기본조사를 위하여 방문한 가구수는 9,290가구이었다. 그리고 가구기본조사를 위하여 방문한 가구 중에서 대상외의 가구(비혈연가구를 제외한 대상가구)는 9,213가구가 되었다. 이 대상가구 중에서 실제로 조사를 완료한 가구는 8,616가구이었다.

이에 관한 세부 내역은 <表 VII-10> 및 <表 VII-11>과 같다. 그리고 가구기본조사 완료가구에서의 개인조사결과는 <表 VII-12>와 같다.

<表 VII-10> 調査結果 家口數

지역	조사구 수	가구수	미방문 가구수	방문결과 가구수		
				계	대상외 가구	대상가구
동	102	6,390	395	5,995	68	5,927
읍·면	56	3,351	56	3,295	9	3,286
계	158	9,741	451	9,290	77	9,213

<表 VII-11> 家口基本調査 結果 家口數 및 人口數

지역	대상 가구수	조사완료 가구수	조사미완 가구수	조사 완료율	조사완료 가구내 인구수
동	5,927	5,479	448	92.4	18,081
읍·면	3,286	3,137	149	95.5	9,411
계	9,213	8,616	597	94.5	27,492

〈表 VII-12〉 個人調査 結果 罹患, 外來利用, 入院利用 調査對象者數

지역	이환조사			외래이용조사			입원이용조사		
	계	완료	미완	계	완료	미완	계	완료	미완
동	6,023	6,006	17	3,755	3,735	20	808	798	10
읍·면	4,080	4,075	5	2,325	2,321	4	514	510	4
계	10,103	10,081	22	6,080	6,056	24	1,322	1,308	14

마. 推定方法

이 조사결과에 의한 각종 통계치는 가중표본합계치(weighted sample total)에 의한 평균 및 구성비이며, 가중표본합계치는 다음과 같이 산출한 것이다.

여기서 실제 조사결과를 그대로 합산한 단순표본합계치(simple sample total)를 사용하지 않은 이유는 각 표본조사구가 표본으로 추출된 확률이 다를 뿐만 아니라 미방문가구율, 적격가구율 및 가구기본조사 조사완료율이 조사구마다 다르기 때문이다.

$$Y' = \sum_h \sum_i M'_{hi} Y_{hi}$$

$$M'_{hi} = M_{hi} (B' / B'')$$

$$M_{hi} = (S_h / n_h S_h) (A_{hi} / A'_{hi}) (B_{hi} / B'_{hi})$$

$$B' = \sum_h \sum_i B'_{hi}$$

$$B'' = \sum_h \sum_i M_{hi} B'_{hi}$$

위의 공식에서

Y' = 어떤 특성 \mathcal{Y} 를 갖는 인구의 추정치

h = 층을 나타내는 첨자

i = 표본조사구를 나타내는 첨자

- M' = 가중표본합계치 산출용 승수
 Y = 특성 y 를 갖는 것으로 조사된 사람수
 M = 가구기본조사 완료가구수를 기준으로 한 총수 추정용 승수
 B = 대상가구수
 B' = 가구기본조사 완료가구수
 S = 크기의 측도
 n = 표본조사구수
 A = 가구수
 A' = 가구기본조사를 위하여 방문한 가구수

이와 같이 산출한 가중표본합계치는 실제 조사결과를 그대로 합산한 단순표본합계치와 가구기본조사가 완료된 총 가구수를 제외하고는 어느 경우에도 모두 다르게 된다. 이에 관한 세부 내역은 <表 VII-13>과 같으며, 가구기본조사가 완료된 가구에서 확인된 개인단위 조사의 대상자수와 이에 관한 가중표본합계치는 <表 VII-14>와 같다.

<表 VII-13> 家口基本調査 結果 加重標本合計値

지 역	조사결과		가중표본합계치	
	가구수 ¹⁾	인구수	가구수	인구수
동	5,479	18,081	6,691	22,081
읍·면	3,137	9,411	1,925	5,775
계	8,616	27,492	8,616	27,856

註: 1) 대상가구 중의 가구기본조사 완료가구수임.

바. 標本誤差

이 보고서에 수록된 각종 통계치는 표본조사에 의하여 얻어진 추정치이므로, 이 표본조사에서 사용한 조사방법과 자료처리방법 등을 그대로 사용하고 조사기간도 같게 하여 모든 조건이 완전히 같은 상태에서 전수조사했다고 가정할 때

〈表 VII-14〉 個人調査 對象者數 및 加重標本合計値

지역	조사결과			가중표본합계치
	계	조사완료	조사미완	
이환조사	10,912	10,888	24	10,745
동	6,665	6,646	19	8,139
읍·면	4,247	4,242	5	2,606
외래이용조사	6,575	6,548	27	6,546
동	4,155	4,133	22	5,061
읍·면	2,420	2,415	5	1,485
입원이용조사	1,427	1,411	16	1,419
동	893	882	11	1,091
읍·면	534	529	5	328

에 얻게 되는 통계치와는 차이가 나게 마련인데, 이 차이는 표본오차때문에 생기는 차이이다. 표본조사결과에 의한 추정치에는 비표본오차와 표본오차의 두가지 형태의 오차가 포함된다.

비표본오차는 조사대상의 누락, 용어정의의 불명확, 질문에 대한 이해의 차이, 일부 응답자의 정확한 답변에 대한 무능력 또는 고의적인 회피, 조사원의 기입착오, 수집된 자료의 부호기입착오 및 입력착오, 조사누락 및 미상의 처리에 의한 오차 등과 같이 조사표의 설계단계부터 자료처리의 완료단계까지 표본추출여부와는 관계없이 복합적인 원인에 의하여 생기는 오차로서, 이 비표본오차는 전수조사에서도 생기게 되는 오차이다.

표본오차는 조사대상이 되는 단위를 전부 조사하지 않고 그 일부인 표본에 대해서만 조사하기 때문에 어쩔 수 없이 생기는 오차로서, 표본조사에서는 어느 경우에도 피할수 없는 오차이며, 전수조사에서는 전혀 생길 수 없는 오차이다.

따라서 표본조사에 의한 추정치와 전수조사했다고 가정할 때의 통계치와의 차이는 표본오차로 설명되어야 하며, 비표본오차와는 관계가 없는 것이다.

그런데 전수조사결과의 통계치는 알 수 없기 때문에 이 차이를 정확하게 안다는 것은 불가능하며, 단지 표본조사로 수집된 자료를 기초로 표본오차의 크기

를 추정하여 확률적으로 설명할 수 있을 뿐이다.

이러한 표본오차의 크기를 나타내는 것으로는 주로 표준오차와 상대표준오차를 사용한다. 표준오차가 갖는 뜻은 표본에서 얻은 추정치와 전수조사했다고 가정할 때의 통계치와의 차이가 표준오차보다 적게되는 확률은 약 68%이며, 표준오차의 2배 이하가 되는 확률은 약 95%라는 뜻이다.

상대표준오차는 추정치와 표준오차를 비교하기 위하여 표준오차를 추정치로 나눈 것으로 일반적으로 백분율로 나타내고 있다.

1) 基本指標에 대한 標本誤差 算出結果

〈表 VII-15〉에는 기본지표에 대한 표본오차를 표준오차와 상대 표준오차로 제시하였으며, 이용에 편리하도록 95% 신뢰구간도 함께 제시하였다.

이 表에서 2주간 외래이용율(100명당)을 예로 들면, 이 표본조사결과의 외래이용율은 23.94%이지만 95% 신뢰구간은 22.97~24.92%이므로 표본조사를 하지 않고 전수조사를 했다고 가정할 때에는 22.97~24.92% 구간 내의 값을 얻을 확률이 약 95%가 된다는 뜻이다.

2) 基本指標에 대한 標本誤差 算出方法

〈表 VII-15〉에서 제시한 비율($R' = Y' / X'$)에 대한 표준오차(se) 및 상대표준오차(rse)는 다음과 같이 산출한 것이다. 여기서 var은 분산을 나타낸다.

$$\text{var}(R') = \sum_h W_h^2 \text{var}(R'_h)$$

$$W_h = X'_h / X'$$

$$\text{var}(R'_h) = \{n_h / 2(n_h - 1) / (X'_h)^2\} \sum_i DZ_{hi}^2$$

$$DZ_{hi} = (Y'_{hi} - Y'_{hi+1}) - R'_h (X'_{hi} - X'_{hi+1})$$

$$\text{se}(R') = \text{var}(R')^{1/2}$$

$$\text{rse}(R') = \text{se}(R') / R' \times 100$$

〈表 VII-15〉 基本指標에 대한 標本誤差¹⁾

항 목	추정치	표준오차	상대표준 오차(%)	95% 신뢰구간
전체 입원(100명당)	5.63	0.21	3.74	5.21~6.05
분만제외 입원	4.36	0.19	4.30	3.99~4.73
연간 만성질환유병률(1,000명당)	431.09	11.17	2.59	408.75~453.43
감염성 및 기생충성	2.51	0.45	17.91	1.61~3.41
신생물	3.22	0.51	15.71	2.20~4.24
내분비장애	20.33	0.89	4.36	18.55~22.11
정신장애	8.07	0.58	7.17	6.91~9.23
신경계 및 감각계	26.68	1.64	6.14	23.40~29.96
순환기계	57.33	2.32	4.04	52.69~61.97
호흡기계	42.07	1.91	4.54	38.25~45.89
소화기계	133.41	4.44	3.33	124.53~142.29
비뇨생식기계	10.54	1.07	10.14	8.40~12.68
피부 및 피하조직	12.13	1.05	8.63	10.03~14.23
근골격계	104.79	3.41	3.25	97.97~111.61
연간 외래이용률(100명당)	67.84	0.86	1.26	66.12~69.55
병·의원	52.05	0.84	1.61	50.37~53.73
치과	21.23	0.65	3.07	19.93~22.53
한방	16.26	0.50	3.06	15.27~17.26
보건기관	18.73	0.90	4.80	16.93~20.53
2주간 급성질환발생률(1,000명당)	131.60	5.22	3.97	121.16~142.04
2주간 외래이용률(100명당)	23.94	0.49	2.03	22.97~24.92
종합병원	2.35	0.13	5.51	2.09~2.61
병원	1.14	0.12	10.85	0.89~1.38
의원	9.33	0.30	3.26	8.72~9.94
치과	1.61	0.13	7.82	1.35~1.86
한방	1.19	0.10	8.12	1.00~1.38
약국	9.75	0.45	4.60	8.85~10.64
보건기관	1.29	0.13	9.74	1.04~1.54

註: 1) 이 表는 예를 들기 위하여 1995년도 보고서의 내용을 그대로 수록한 것임.

3. 1998年度 保健意識行態調査 標本設計概要(案)

가. 企劃段階에서의 決定事項

1) 調査對象地域

각 시·도의 섬지역을 제외한 전국을 조사대상지역으로 하였다. 표본조사구는 전국을 대상지역으로 하여 추출하는 것이 바람직하지만, 교통이 불편한 섬지역에서 표본조사구가 추출되면 현실적으로 조사가 곤란하게 되는 경우가 있으므로 각 시·도의 섬지역은 조사대상지역에서 제외시켰다(제외된 섬지역의 총 가구수는 전국 가구수의 0.6% 미만임).

2) 調査對象家口

조사대상가구에는 보통가구와 1인가구를 포함시켰다. 여기서 보통가구라 함은 원칙적으로 혈연관계가 있는 2인 이상의 사람들이 생활을 같이하고 있는 경우로서 일반적으로 말하는 가구의 개념이며, 비혈연 가구원이 함께 있는 경우도 있다.

그리고 1인가구는 혼자서 자취를 하는 경우와 다방, 상점, 음식점, 미장원 등에서 종업원 1명이 혼자서 숙식을 하는 경우 등을 말한다.

일반적인 가구조사의 경우에는 친구와 함께 자취를 하거나 다방, 상점, 음식점, 미장원 등에서 혈연관계가 없는 종업원 2인 이상이 함께 숙식을 하는 경우와 같은 비혈연가구도 포함되나, 국민건강조사에서는 비혈연가구의 경우에는 가구원 서로가 다른 사람의 의료이용, 이환 등에 관한 사항을 알지 못하는 경우가 대부분이기 때문에 현실적으로 조사가 불가능하므로 조사대상 가구에서 제외시켰다.

이와 같은 이유로 보통가구의 가구원 중에서도 비혈연 가구원은 조사대상 가구원에서 제외시켰다.

보건·의식행태조사는 국민건강조사의 조사대상 가구원 중에서 15~69세의 가구원을 조사대상 가구원으로 하였다.

3) 標本規模

조사규모는 1995년도 조사의 조사구수(110개 조사구)와 가구수(7,700가구)보다는 크게 되도록 하되, 주어진 예산과 조사관리능력 등을 고려하여 조사구수는 160개 조사구를, 그리고 가구수는 10,000가구를 초과하지 않도록 하였다.

표본조사구는 1997~2001년까지 5년간 각종 가구단위 통계조사에 다목적으로 사용하기 위하여, 1995년 인구총조사 조사구 중에서 각 시·도의 섬지역 조사구와 시설단위 조사구를 제외한 보통조사구를 이용하여 1996년도에 추출한 기본표본조사구 중에서 추출기로 하였다.

기본표본조사구의 추출에 사용된 ‘인구주택총조사의 보통조사구’를 편의상 다음부터는 ‘인구총조사구’란 용어로 바꾸어 사용하기로 한다.

기본표본조사구는 150개 기본표본조사구의 제1표본, 200개 기본표본조사구의 제2표본, 250개 기본표본조사구의 제3표본으로 편성되어 있으며, 각 기본표본조사구는 5~7개(평균 5.6개)의 인구총조사구가 통합된 것이다. 이러한 기본표본조사구의 평균 가구수는 인구총조사의 가구수로 350가구이며, 기본표본조사구에 포함된 인구총조사구의 평균 가구수는 63가구이다.

실사표본조사구는 160개 조사구를 초과할 수 없으므로 제1표본의 150개 기본표본조사구에서 각각 1개씩의 인구총조사구를 추출하여 실사표본조사구로 하면 된다.

그런데 표본으로 추출된 인구총조사구 내에 1995년 인구총조사 이후에 주택이 많이 신축되었을 때(특히 아파트가 신축되었을 때)에는 가구수가 많이 증가되어도 경계선 안의 모든 가구를 전부 조사하여야 하지만, 가구수가 너무 많이 증가한 경우에는 모든 가구를 전부 조사하는 것이 현실적으로 곤란하게 되는 경우가 있어, 자료수집이 가능한 1995년 인구총조사 이후의 신축 아파트에 대하여는 별도로 표본조사구를 추출하여 추가기로 하였다.

인구총조사구의 총 가구수는 12,885,650가구이며, 1995년 인구총조사(1995. 11. 1) 이후 1997년 10월말까지의 신축 아파트 가구수는 약 62,000가구분으로 예상되므로 가구수의 비는 100.0:4.8이 된다.

따라서 이들 두 자료에서 같은 비율로 실사표본조사구를 추출한다면 인구총

조사구에서 150개의 조사구를 추출하게 되므로 신축 아파트에서는 7개($4.8 \times 150 / 100 = 7.2$)조사구를 추출하여야 한다.

결과적으로 실사표본조사구는 기본표본조사구의 제1표본의 150개 기본표본조사구에서 각각 1개씩의 인구총조사구를 실사표본조사구로 추출하고 신축 아파트에서는 60가구 기준의 7개 실사표본조사구를 추출키로 하였다(실제적으로는 신축 아파트의 조사구 추출과정에서 8개 조사구로 조정됨).

이와 같이 157개의 실사표본조사구를 추출하면 자료상의 조사구 내의 가구수(150개 인구총조사구의 가구수 + 7개 신축 아파트조사구의 가구수)는 약 9,600가구가 될 것이며, 1995년 인구총조사 이후의 가구증가를 고려하면 실제 조사 때에는 9,700가구가 될 것으로 보았다. 그리고 불응, 장기 출타 및 기타 사유에 의한 조사미완률을 4%로 보면 9,310가구에 대하여 가구기본조사를 완료하게 되며, 1995년도 조사결과의 대상외의 가구율(비혈연가구율) 0.8%를 고려하면 대상 가구는 9,240가구가 될 것으로 기대하였다.

대상가구 중에서 대상자(15~69세)가 없는 부적격 가구율을 4.9%(1995년도 조사의 경우 4.6%)로 보면 적격가구는 8,790가구가 되며, 조사대상자는 각 가구에서 생일(년월 무시)이 가장 빠른 사람을 추출하면 총 8,790명이 될 것으로 기대하였다.

나. 基本標本調査區 抽出概要

이 부분은 '1997년도 출산력 및 가족보건실태조사 표본설계개요(안)'에서 설명한 '기본표본조사구 추출개요'와 같으므로 여기서는 생략한다.

다. 實查標本調査區의 抽出

이 부분은 '1998년도 국민건강조사 표본설계개요(안)'에서 설명한 '실사표본조사구의 추출'과 같으므로 여기서는 생략한다.

라. 調査結果 家口數 및 人口數

표본조사구에서는 모든 가구에 대하여 전수조사토록 하였으며, 이를 위하여 지

도원은 3명의 조사원에게 각 조사원이 담당하여야 할 구역을 설정하여 주었다.

각 조사원은 담당 구역 내의 가구를 거쳐 단위의 순서로 방문하면서 가구명부에 각 가구에 관한 사항을 기입한 다음, 가구기본조사표를 먼저 작성하고, 이에 따라 해당 개인조사표를 작성토록 하였다.

그런데, 주택의 신축으로 가구가 많이 증가한 조사구, 낮에는 응답자를 만날 수 없는 경우가 특히 많은 조사구, 집단불응으로 설득에 많은 시간이 소요되는 조사구에서는 모든 가구를 전부 조사하려면 조사기간을 연장하여야 한다.

그러나 조사기간 연장에 수반되는 예산과 다음 조사구가 속한 모든 동·읍·면에 조사기간 변경내용을 통보하여야 하는 등의 행정절차가 복잡하게 된다. 따라서 각 조사구별로 정해진 기간 내에 조사를 못하게 된 가구는 연구책임자의 허락을 받아 가구명부만 작성토록 하였다.

현지조사 결과, 158개 표본조사구의 가구명부에 기입된 총가구수는 9,741가구가 되었으며, 이 중에서 가구기본조사를 위하여 방문한 가구수는 9,290가구이었다.

그리고 가구기본조사를 위하여 방문한 가구 중에서 대상외의 가구(비혈연가구를 제외한 대상가구는 99.2%에 해당하는 9,213가구가 되었으며, 이에 관한 지역별 내역은 <表 VII-16>과 같다.

<表 VII-16> 調査結果 家口數

지역	조사구수	가구수	미방문 가구수	방문결과 가구수		
				계	대상외 가구	대상가구(%)
동	102	6,390	395	5,995	68	5,927(98.9)
읍·면	56	3,351	56	3,295	9	3,286(99.7)
계	158	9,741	451	9,290	77	9,213(99.2)

조사대상이 된 9,213가구 중에서 실제로 조사를 완료한 가구는 8,616가구이였으며, 이 중에서 15~69세의 사람이 없는 부적격가구(378가구)를 제외한 적격가구는 8,238가구가 되었다.

적격가구에서는 각각 15~69세의 가족원 중에서 생일이 가장 빠른 사람을 보건 의식행태조사의 조사대상자로 하였다.

따라서 조사대상자는 8,238명이 되었으며, 이 중에서 7,502명(91.1%)에 대하여 조사를 완료하였다.

이에 관한 지역별 내역은 <表 VII-17> 및 <表 VII-18>과 같다.

<表 VII-17> 家口基本調査 結果 家口數

지역	조사구수	방문가구 중의 대상가구수	조사완료 가구수	조사미완 가구수	조사 완료율
동	102	5,927	5,479	448	92.4
읍·면	56	3,286	3,137	149	95.5
계	158	9,213	8,616	597	94.5

<表 VII-18> 保健意識行態調査 結果 對象者數

지역	가구기본조사 완료가구수	보건 의식행태조사			가중표본 합계치
		대상자수	조사완료	조사미완 조사완료율	
동	5,479	5,365	4,770	595 88.9	6,512
읍·면	3,137	2,873	2,732	141 95.1	1,726
계	8,616	8,238	7,502	736 91.1	8,238

마. 推定方法

이 조사결과에 의한 각종 통계치는 다음에 제시한 방법으로 산출한 가중표본합계치(weighted sample total)와 이를 기초로 한 평균 및 구성비이다.

여기서 단순표본합계치(simple sample total)를 사용하지 않고 가중표본합계치를 사용한 이유는 각 표본조사구가 표본으로 추출된 확률이 동일하지 않고 조사완료율이 표본조사구마다 다르기 때문이다.

이 조사결과의 가중표본합계치는 다음과 같이 산출한 것이다.

$$Y' = \sum_h \sum_i \sum_j M'_{hij} Y_{hij}$$

$$M'_{hij} = M''_{hij} (C/P'')$$

$$C = \sum_h \sum_i C_{hi}$$

$$P'' = \sum_h \sum_i \sum_j M''_{hij} P_{hij}$$

$$M''_{hij} = M_{hi} (C_{hi}/C'_{hi})$$

$$M_{hi} = (S_{hi}/n_h S_{hi}) (A_{hi}/A'_{hi}) (B_{hi}/B'_{hi})$$

위의 공식에서

Y' = 어떤 특성 γ 를 갖는 인구의 추정치

h = 층을 나타내는 첨자

i = 표본조사구를 나타내는 첨자

j = 가구를 나타내는 첨자

M' = 가중표본합계치 산출용 승수

M'' = 총수추정용 승수

Y = 특성 γ 를 갖는 것으로 조사된 인구수

M = 가구기본조사 완료가구수를 기준으로 한 총수 추정용 승수

B = 적격가구수

B' = 가구기본조사 완료가구수

S = 크기의 측도

n = 표본조사구수

A = 가구수

A' = 가구기본조사를 위하여 방문한 가구수

C = 표본으로 추출된 인구수

C' = 표본으로 추출된 인구 중에서 조사완료 인구수

P = 조사대상 인구수

P'' = 조사모집단의 조사대상 총인구수 추정치

4. 2000年度 障碍人實態調査 標本設計概要(案)

가. 企劃段階에서의 決定事項

1) 調査母集團

가) 調査對象地域

이 부분은 '1997년도 출산력 및 가족보건실태조사 표본설계(안)'의 조사대상 지역과 같으므로 여기서는 생략한다.

나) 調査對象家口 및 家口員

보통가구, 1인가구 및 5인 이하의 비혈연가구인 일반가구를 조사대상가구로 하였으며, 조사대상 가구에서 통상적으로 생활을 같이 하는 사람을 조사대상 가구원으로 하였다.

여기서 보통가구나 함은 원칙적으로 혈연관계가 있는 2인 이상이 생활을 같이 하고 있는 경우로서 일반적으로 말하는 가구의 개념이며, 비혈연 가구원이 함께 있는 경우도 있다. 1인가구는 혼자서 자취를 하는 경우와 다방, 상점, 음식점, 미장원 등에서 종업원 1명이 숙식을 하는 경우를 말하며, 비혈연가구는 혈연관계가 없는 5명 이하의 사람들이 함께 모여 가구의 형태를 유지하면서 사는 경우(예: 친구와 함께 자취를 하는 경우와 다방, 음식점, 기숙사가 아닌 공장 내의 방에서 전혀 혈연관계가 없는 5명 이하의 종업원이 함께 숙식을 하는 경우)를 말한다.

따라서 6인 이상의 비혈연가구와 기숙사 및 특수사회시설(아동복지법, 노인복지법, 장애인복지법에 의하여 설치된 수용시설, 특수병원, 사설 기도원 및 이에 준하는 시설)은 집단가구로 취급하여 조사대상가구에서 제외하였다.

보통가구 내에 혈연관계의 가구원 외에 비혈연관계의 사람도 함께 생활하고 있는 경우에는 3명까지는 보통가구의 가구원으로 취급하고 4~5명인 경우에는 별도의 비혈연가구로 분리시키기로 하였다.

2) 標本規模

조사규모는 1995년 조사의 조사구수(200개 조사구)와 가구수(41,300가구)보다는 크게 되도록 하되, 주어진 예산과 조사관리 능력을 종합적으로 고려하여 조사구수는 250개 조사구를 초과하지 않도록 하고, 가구수는 52,000가구를 초과하지 않도록 하였다.

표본조사구는 1997~2001년까지 5년간 각종 가구단위 통계조사에 다목적으로 사용하기 위하여, 1995년 인구주택총조사 조사구 중에서 각 시·도의 섬지역 조사구와 시설단위 조사구를 제외한 보통조사구를 이용하여 1996년도에 추출한 기본표본조사구를 이용하여 추출기로 하였다.

기본표본조사구의 추출에 사용된 '인구주택총조사의 보통조사구'를 편의상 다음부터는 '인구총조사구'란 용어로 바꾸어 사용하기로 한다.

기본표본조사구는 150개 기본표본조사구의 제1표본, 200개 기본표본조사구의 제2표본 및 250개 기본표본조사구의 제3표본으로 편성되어 있으며, 기본표본조사구는 각각 5~7개(평균 5.6개)의 인구총조사구가 통합된 것이다. 기본표본조사구의 평균 가구수는 인구총조사의 가구수로 350가구이며, 기본표본조사구에 포함된 인구총조사구의 평균 가구수는 63가구이다.

장애인실태조사는 장애인수 추정이 1차적인 목적이며, 인구 100명당 장애인수는 1995년도 조사결과 2.35명 정도의 발생빈도가 낮은 경우를 찾아내어야 하는 조사이므로 실사표본조사구의 가구수는 일반적인 가구조사와는 달리 적어도 평균 200가구 이상이 되는 것이 바람직하다.

기본표본조사구의 인구총조사구당 평균 가구수는 63가구이므로 200가구 이상이 되도록 3개 또는 4개의 인구총조사구를 통합하여 실사표본조사구로 하면 실사표본조사구의 평균 인구총조사구수는 3.5개가 되며 평균 가구수는 220가구가 된다. 이 때 조사가구수를 52,000가구에 접근하도록 하려면 236개 실사표본조사구를 추출하여야 한다($220\text{가구} \times 236\text{개 조사구} = 51,920\text{가구}$).

그런데 표본으로 추출된 조사구 경계선 안에 1990년 인구총조사 이후에 주택이 많이 신축되었을 때(특히 아파트가 신축되었을 때)에는 가구수가 많이 증가되어도 경계선 안의 모든 가구를 다 조사하여야 하지만, 가구수가 너무 많이 증가

한 경우에는 모든 가구를 전부 조사하는 것이 현실적으로 곤란하게 되는 경우가 있을 뿐만 아니라, 표본오차도 증가하게 된다.

이에 대한 대책으로 조사구 내의 가구 중에서 일부 가구를 표본으로 추출하여 조사하고 추정과정에서 이에 따르는 가중치를 적용하는 방법도 있지만, 현지 조사팀이 연구진의 지시를 받아 표본가구를 선정하려면 표본가구 추출을 위한 가구명부를 작성하여야 하는 등의 현지업무가 복잡하게 되며, 표본오차가 증가하는 것은 마찬가지이다. 이러한 부담을 조금이라도 감소시키기 위하여 일부의 실사 표본조사구는 자료수집이 가능한 1990년 인구총조사 이후의 신축 아파트에서 추출하기로 하였다.

그러나 기존 주택을 다세대주택 또는 다가구주택으로 개축한 경우가 많은 조사구와 공지에 아파트가 아닌 주택이 많이 신축된 조사구의 경우에는 1995년 인구총조사 당시보다 가구수가 많이 증가되어 조사기간을 연장하여야 하는 경우가 발생할 수도 있으므로, 이에 대한 추가소요예산을 준비해 두어야 한다.

따라서 실사표본조사구수는 236개 조사구에서 232개 조사구로 축소조정하였다. 232개 조사구의 경우 51,000가구의 조사를 목표로 하게 되며(220가구 × 232개 조사구 = 51,040가구), 조사완료율을 95%로 보면 조사완료가구수는 48,450가구가 될 것으로 보았다.

나. 基本標本調査區 抽出概要

이 부분은 '1997년도 출산력 및 가족보건실태조사 표본설계개요(안)'에서 설명한 '기본표본조사구 추출개요'와 같으므로 여기서는 생략한다.

다. 實查標本調査區의 抽出

인구총조사구의 총 가구수는 12,885,650가구이며, 1995년 인구총조사(1995. 11. 1) 이후 1999년 10월말까지의 신축 아파트 자료를 수집하여 집계한 결과의 전체 규모는 1,238,764가구분이므로, 가구수의 비는 100.0:9.6이 된다.

따라서 232개 실사표본조사구를 이 비율이 되도록 하려면 인구총조사구 부문에서 212개 조사구와 신축아파트 부문에서 20개 조사구를 추출하여야 한다.

1) 基本標本調査區를 利用한 實查標本調査區의 抽出

제3기본표본의 250개 기본표본조사구 중에서 212개의 기본표본조사구를 추출하기 위하여 우선 600개 전체 기본표본조사구의 층별 분포에 비례하도록 212개의 조사구를 각 층에 배분하여 반올림하였으며, 그 결과는 <表 VII-19>의 배분결과와 같다.

<表 VII-19> 212個 調査區의 層別 配分結果

층	전체 기본표본조사구	제3기본표본조사구	배분결과
(1) 6대 시의 동	231 (38.50)	96	82
(2) 기타 시의 동	155 (25.83)	65	55
(3) 시의 읍·면	86 (14.33)	36	30
(4) 군의 읍	46 (7.67)	19	16
(5) 군의 면	82 (13.67)	34	29
합계	600(100.00)	250	212

그리고 각층별로 제3기본표본조사구 중에서 배분된 수의 기본표본조사구를 계통추출하였으며, 추출된 기본표본조사구에서는 각각 200가구 이상이 되도록 3~4개의 인구총조사구를 통합하여 실사표본조사구로 하였다. 이때의 실사표본조사구의 크기의 측도는 각 인구총조사구의 크기의 측도의 합계로 하였다.

이와 같이 추출한 212개 실사표본조사구의 인구총조사 가구수는 46,852가구가 되었으며, 크기의 측도는 4,695가 되었다.

2) 新築 아파트의 實查標本調査區 抽出

1,238,764가구분의 신축아파트를 6대 시의 동(제1층), (2) 기타 시의 동(제2층), (3) 읍·면(제3층)의 3개 층으로 층화하였다.

그리고 인구총조사구에서 기본표본조사구를 추출할 때와 같이 동(제1층 및 제2층)의 가구수는 그대로 두고 읍·면(제3층)의 가구수는 2배를 하여 20개의 조사구를 각 층에 배분하였으며, 그 결과는 <表 VII-20>과 같다.

〈表 VII-20〉 層別 新築 아파트의 實查標本調査區數 決定 內譯

층	가구수	크기의 측도	조사구수
(1) 6대 시의 동	503,064 (40.61)	50,306	7.08 → 7
(2) 기타 시의 동	553,848 (44.71)	55,485	7.80 → 8
(3) 읍·면	181,852 (14.68)	18,185	5.12 → 5
계	1,238,764(100.00)	123,976	20

표본조사구의 추출을 위하여 제1층(6대 시의 동)에서는 신축 아파트를 행정구역번호에 따라 서울, 부산, 대구, 인천, 광주, 대전의 순서로 분류하였으며, 제2층(기타 시의 동)의 경우에는 경기, 강원, ..., 제주의 순서로 분류하였다. 제3층(읍·면)의 경우에는 6대 시의 읍, 기타 시의 읍, 6대 시의 면, 기타 시의 면, 군의 읍, 군의 면 순서로 먼저 분류하고 각 분류 내에서는 서울, 부산, ..., 대전, 경기, ..., 제주의 순서로 재분류하였다. 각 층에서 이와 같이 신축아파트를 1차 분류한 다음, 시·도내에서는 다시 신축 아파트 소재지의 구·시·군 명칭의 '가나다' 순서로 정리하였다.

그리고 제1층(6대 시의 동)에서는 7개의 아파트를, 제2층(기타 시의 동)에서는 8개의 아파트를, 제3층(읍면)에서는 5개의 아파트를 가구수에 의한 확률비례계통추출법으로 추출하였다. 표본으로 추출된 각 아파트에서는 첫 번째 동의 첫 번째 계단(복도식인 경우에는 1층)부터 순서대로 220가구분이 포함되는 조사구를 설정하여 표본조사구로 하였다.

따라서 1개 표본조사구에는 1개 동 이상의 아파트가 포함되었으며, 그 중 마지막 동의 경우에는 일부만 포함된 경우가 있게 되었다. 이때의 표본조사구의 크기의 측도는 22로 하였다.

결과적으로 신축 아파트에서 추출된 20개 조사구의 가구수는 4,400가구가 되었으며, 전체 크기의 측도는 440이 되었다.

3) 實查標本調査區 抽出結果의 標本規模

이와 같이 232개의 실사표본조사구를 추출한 결과의 총 가구수는 51,290가구

가 되었으며, 그 세부내역은 <表 VII-21>과 같다.

<表 VII-21> 實查標本調査區 抽出結果의 標本規模

자료 및 층	모집단의 크기의 측도	실사표본			표본비율
		조사구수	크기의 측도	가구수 ¹⁾	
인구총조사구					
(1) 6대 시의 동	608,395	82	1,808	18,042	1/337
(2) 기타 시의 동	408,580	55	1,217	12,132	1/336
(3) 시의 읍·면	112,966	30	662	6,588	1/171
(4) 군의 읍	61,178	16	355	3,525	1/172
(5) 군의 면	108,048	29	640	6,354	1/169
소계	1,299,167	212	4,682	46,641	1/277
신축 아파트					
(6) 6대 시의 동	50,306	7	154	1,540	1/327
(7) 기타 시의 동	55,485	8	176	1,760	1/315
(8) 읍·면	18,185	5	110	1,100	1/165
소계	123,976	20	440	4,400	1/282
동	1,122,766	152	3,355	33,474	1/335
읍·면	300,377	80	1,767	17,567	1/170
계	1,423,143	232	5,122	51,041	1/278

註: 1) 인구총조사구의 경우 인구주택총조사 결과의 일반가구수이며, 신축아파트의 경우 준공검사결과와 가구수(호수)임.

라. 調査結果 家口數 및 人口數

232개 실사표본조사구에서는 모든 가구에 대하여 전수조사하였다. 조사결과와 총가구수는 기초자료의 가구수보다 670가구가 증가한 51,711가구가 되었으며, 이 중에서 94.9%에 해당하는 49,096가구에 대하여 조사를 완료하였다. 그리고 조사완료된 가구의 총인구는 151,449명이 되었다. 이에 관한 세부 내역은 <表 VII-22>와 같다.

〈表 VII-22〉 家口基本調査 結果 家口數

지역	조사구수	가 구 수			조사 완료율	조사완료 인구수	가구내의 장애인수
		계	조사완료	조사미완			
동	152	33,950	32,082	1,868	94.5	103,946	2,050
읍·면	80	17,390	16,607	783	95.5	48,326	2,426
계	232	51,711	48,689	2,651	94.9	152,272	4,476

마. 推定方法

1) 總數의 推定值

이 조사결과에서 어떤 특성 γ 를 갖는 인구의 추정치는 다음과 같이 산출하였다.

$$Y' = \sum_h Y'_h = \sum_h \sum_i Y'_{hi} = \sum_h \sum_i M'_{hi} Y_{hi} \quad \dots (1)$$

$$M'_{hi} = M_{hi} (P''/P') \quad \dots (2)$$

$$M_{hi} = (S_h/n_h S_{hi}) (A'_{hi}/A_{hi}) \quad \dots (3)$$

$$P' = \sum_h \sum_i M_{hi} P_{hi} \quad \dots (4)$$

위의 공식에서

Y = 어떤 특성 γ 를 갖는 인구의 추정치

h = 층을 나타내는 첨자

i = 표본조사구를 나타내는 첨자

M' = 추정치 산출용 조정 승수

Y = 특성 γ 를 갖는 것으로 조사된 사람수

M = 제1단계 승수

P'' = 가구조사 모집단의 기준 인구(45,947,856명)

P' = 가구조사 모집단의 제1단계 추정인구(45,413,831명)

S = 크기의 측도

n = 표본조사구수

A' = 표본조사구 내의 총가구수(조사완료+조사미완)

A = 조사완료 가구수

P = 조사된 인구수

공식 (2)에서, P'' 는 조사기간 중간 시점인 2000년 4월 15일 현재의 가구조사 모집단의 기준 인구이며, 이는 다음과 같이 결정한 것이다.

장래인구추계(통계청, 1996)에 의하면 1999년 7월 1일과 2000년 7월 1일 현재의 인구는 각각 46,858,463명과 47,274,543명이므로 1년간 416,080명이 증가한 결과가 된다. 따라서 월평균 증가 인구는 34,673명이므로, 2000년 4월 15일까지 9.5개월 동안의 증가 인구는 329,393명이 된다. 이에 따라 2000년 4월 15일 현재의 인구는 47,187,856명으로 하였다.

이번 조사에서는 집단가구(6인 이상의 일반 비혈연가구와 집단 거주 및 수용 시설의 비혈연가구)가 제외되었으므로, 가구조사 모집단의 인구는 2000년 4월 15일 현재 인구에서 집단가구의 인구를 제외하여야 한다.

1995년 인구총조사 결과에 의하면 총인구가 44,551,183명이고 일반가구의 인구는 43,319,778명이므로 집단가구의 인구는 1,231,405명이 된다. 따라서 이를 근거로 2000년 4월 15일 현재의 집단가구의 인구를 약 1,240,000명으로 보고, 2000년 4월 15일 현재 인구 47,187,856명에서 이 집단가구의 인구 1,240,000명을 뺀 45,947,856명을 조사모집단의 기준 인구로 택하였다.

그런데 공식 (4)에 의하여 산출한 제1단계 추정인구는 45,413,832명이며, 이는 기준 인구보다 534,024명이 작다.

여기서 조사대상지역에서 제외된 섬지역의 인구를 약 250,000명으로 보면, 실제적으로는 284,025명이 적은 결과가 된다.

따라서, 섬지역에서의 장애출현을 및 분포가 조사모집단의 경우와 같다는 가정하고 제1단계 추정인구를 기준 인구가 되도록 하기 위하여 공식 (3)의 제1단계 승수를 공식 (2)와 같이 조정하여 적용하였다.

2) 構成比, 比率 및 平均의 推定值

특성 x 에 대한 특성 y 의 구성비(비율) 및 평균의 추정치는 다음과 같이 산출하였다.

$$R' = Y' / X' \quad \dots\dots (5)$$

위의 공식에서

R' = 특성 x 에 대한 특성 y 의 구성비, 비율 및 평균의 추정치

Y' = 특성 y 에 대한 인구의 총수 추정치

X' = 특성 x 에 대한 인구의 총수 추정치

3) 加重標本合計值

특성 y 를 갖는 인구에 대한 가중표본합계치는 조사된 장애인수를 기준으로 다음과 같이 산출하였다.

$$y' = \sum_h Y'_h = \sum_h \sum_i Y'_{hi} = \sum_h \sum_i M''_{hi} Y_{hi} \quad \dots\dots (6)$$

$$M''_{hi} = M'_{hi} (Z/Z') \quad \dots\dots (7)$$

위의 공식에서

y' = 특성 y 를 갖는 인구에 대한 가중표본합계치

M'' = 가중표본합계치 산출용 승수

Z = 표본에서 조사된 장애인수(4,476명)

Z' = 공식 (1)에서 추정된 장애인수(1,171,448명)

바. 標本誤差

이 보고서에 수록된 각종 통계치는 표본조사에 의하여 얻어진 추정치이므로, 이 표본조사에서 사용한 조사방법과 자료처리방법 등을 그대로 사용하고 조사기간도 같게 하여 모든 조건이 완전히 같은 상태에서 전수조사했다고 가정할 때에 얻게되는 통계치와는 차이가 나게 마련이다.

표본조사결과에 의한 추정치에는 비표본오차와 표본오차의 두가지 형태의 오차가 포함된다.

비표본오차는 조사대상의 누락, 응답자의 답변착오, 조사원의 기입착오, 수집된 자료의 부호기입착오 및 입력착오 등과 같이 조사표의 설계단계부터 자료처리의 완료단계까지 표본추출 여부와는 관계없이 복합적인 원인에 의하여 생기는 오차로서, 이 비표본오차는 전수조사에서도 생기게 되는 오차이다.

표본오차는 조사대상이 되는 단위를 전부 조사하지 않고 그 일부인 표본에 대해서만 조사하기 때문에 어쩔 수 없이 생기는 오차로서, 표본조사에서는 어느 경우에도 피할수 없는 오차이며, 전수조사에서는 전혀 생길 수 없는 오차이다.

따라서 표본조사에 의한 추정치와 전수조사했다고 가정할 때의 통계치의 차이는 표본오차로 설명되어야 한다. 그런데 전수조사결과의 통계치는 알 수 없기 때문에 이 차이를 정확하게 안다는 것은 불가능하며, 단지 표본조사로 수집된 자료를 기초로 표본오차의 크기를 추정하여 확률적으로 설명할 수 있을 뿐이다.

이러한 표본오차의 크기를 나타내는 것으로는 주로 표준오차를 사용하는데, 표준오차가 갖는 뜻은 표본에서 얻은 추정치와 전수조사했다고 가정할 때의 통계치의 차이가 표준오차보다 적게되는 확률은 약 68%이며, 표준오차의 2배 이하가 되는 확률은 약 95%라는 뜻이다.

이번 조사결과의 기본 항목에 대한 표준오차 산출결과는 <表 VII-23>과 같다.

<表 VII-23> 在家 障碍人數 및 出現率의 標準誤差

장애종류	장애인수(명)		출현률(1,000명당)	
	추정치 ¹⁾	표준오차	추정치	표준오차
전 체	1,171,448	33,562	25.50	0.80
지체장애	792,759	27,077	17.25	0.64
시각장애	83,237	6,191	1.81	0.15
청각장애	174,713	8,688	3.80	0.20
언어장애	41,413	3,734	0.90	0.08
정신지체	79,326	4,809	1.73	0.11

註: 1) 중복장애의 경우 주된 장애로 분류

사. 標準誤差 算出公式

1) 總數의 標本誤差

공식 (1)에 의하여 산출한 Y' 에 대한 분산(var) 및 표준오차(se)의 추정식은 다음과 같다.

$$\text{var}(Y') = \sum_h \{n_h / 2 (n_h - 1)\} \sum_i (DY'_{hi})^2 \quad \dots (8)$$

$$DY'_{hi} = Y'_{hi} - Y'_{hi+1} \quad \dots (9)$$

$$\text{se}(Y') = \text{var}(Y')^{1/2} \quad \dots (10)$$

$$\text{var}(Y') = \sum_h \{n_h / 2 (n_h - 1)\} \sum_i (DY'_{hi})^2 \quad \dots (11)$$

2) 出現率 및 構成比의 標本誤差

공식 (5)에 의하여 산출한 R' 에 대한 분산(var), 표준오차(se)의 추정식은 다음과 같다.

$$\text{var}(R') = \sum_h W_h^2 \text{var}(R'_h) \quad \dots (12)$$

$$W_h = X'_h / X \quad \dots (13)$$

$$\text{var}(R'_h) = \{n_h / 2 (n_h - 1) / (X'_h)^2\} \sum_i (DZ_{hi})^2 \quad \dots (14)$$

$$DZ_{hi} = DY'_{hi} - R'_h DX'_{hi} \quad \dots (15)$$

$$\text{se}(R') = \text{var}(R')^{1/2} \quad \dots (16)$$

VIII. 結 論

본 연구에서는 1997년부터 2001년까지 5년간 각종 가구단위통계조사의 실사 표본조사구 추출에 사용하기 위하여 1995년 인구주택총조사의 조사구 중에서 시설단위의 조사구와 각 시·도의 섬지역 조사구를 제외한 219,771개 보통조사구 중에서 3,363개 인구주택총조사 조사구로 구성되는 600개의 기본표본조사구를 추출하였다.

각 기본표본조사구에는 5~7개의 인구주택총조사 조사구(평균 5.6개 조사구)가 포함되었으며, 인구주택총조사 가구수로는 320가구 이상이 되었다.

이 600개의 기본표본조사구는 제1표본(150개의 기본표본조사구), 제2표본(200개의 기본표본조사구), 제3표본(250개의 기본표본조사구)으로 편제하였다.

기본표본조사구 중에서 추출하는 실사표본조사구가 150개 조사구 이하인 경우에는 제1표본을 사용하고 151~200개 조사구인 경우에는 제2표본을, 그리고 201~250개 조사구인 경우에는 제3표본을 사용하면 될 것이다.

한편, 첫 번째 조사에서 제1표본을 사용하였는데, 다음 번의 조사에서도 실사 표본조사구가 150개 조사구 이하인 경우에는 제1표본을 사용하지 않고 제2표본을 사용하면 된다. 마찬가지로 어떤 조사에서 제2표본을 사용하였는데 다음 번의 다른 조사에서도 실사표본조사구가 150~200개 조사구인 경우에는 제3표본을 사용하면 된다. 이러한 방법을 택하면, 동일한 동·읍·면이 2번 연속하여 해당되는 것을 방지할 수 있게 된다.

만일 어떤 조사에서 제3표본의 250개 기본표본조사구보다 많은 수의 실사표본조사구가 필요하게 된다면 제1표본, 제2표본, 제3표본의 3개 표본 중에서 2개의 표본을 통합하여 사용하면 될 것이다.

본 연구에서 추출한 기본표본조사구를 가구단위 통계조사에 사용하기 위해서는 실사표본조사구를 추출하여야 하며, 이를 토대로 현지조사를 실시하다 보면 여러 가지의 애로점이 생기게 된다. 따라서 기본표본조사구의 유지관리 방법과 유의사항을 언급함으로써, 결론에 대신하고자 한다.

1. 基本標本調査區의 維持·管理方法

가. 基本標本調査區의 行政區域名 修正·補完

구·시·군 또는 행정 동·읍·면의 신설, 통폐합 및 관할 구역 변경 등의 행정구역개편은 주민을 위한 행정의 효율적인 수행을 위하여 법령 또는 구·시·군 조례에 의하여 수시로 실시되고 있다. 특히 행정 동·읍·면의 경우에는 명칭만 변경되는 경우도 있다.

기본표본조사구 명부의 행정구역명은 1996년 6월 30일 현재의 행정구역명이므로 그 이후의 행정구역 변경에 따라 주기적으로 수정·보완하여야 한다. 만일 행정구역명을 수정·보완하지 않으면 실제적으로는 A동 관할구역의 조사구이나 B동 관할구역의 조사구가 되므로 실사단계에서 관련 공문이 B동사무소로 가게 되어 혼선을 빚게 되며, 이로 인하여 실사계획에 차질이 생기게 된다.

행정구역개편에 관한 자료는 통계청(통계기획국 통계기준과)에서 분기별로 수집하여 8개 종류로 나누어 정리하고 있으며, 그 내용은 다음과 같다.

- ① 구(자치구 및 비자치구)의 신설 및 분구 내역
- ② 도농 복합형태의 통합시 신설 내역
- ③ 읍승격(도농 복합시 지역) 내역
- ④ 행정동 신설 및 분동 내역
- ⑤ 행정동 통합내역
- ⑥ 구·시·군 및 동·읍·면 명칭변경
- ⑦ 행정 관할구역 변경 내역
- ⑧ 시·도 출장소 현황

이와 같은 자료는 담당자에게 의뢰하면 제공받을 수 있으므로, 매 분기별로 제공받아 처리토록 하여야 한다. 따라서 매 분기별로 행정구역 개편지역 중에서 기본표본조사구와 관계되는 행정구역이 있는지를 확인하여, 기본표본조사구가 속한 행정구역이 개편되었을 때에는 기본표본조사구의 행정구역명을 수정·보완하여야 한다.

2개의 행정동이 1개의 행정동으로 통합된 경우에는 그 처리가 간단하게 된다.

그러나 1개의 행정동이 2개의 행정동으로 분동된 경우, 관할구역의 일부가 다른 행정동으로 편입된 경우, 2개의 행정동이 3개의 행정동으로 통폐합된 경우, 읍 또는 면의 일부가 인접 동·읍·면으로 편입된 경우에는 기본표본조사구에 속한 각 인구총조사구가 실제로 어느 행정동에 속하는지를 관련 동·읍·면사무소에 전화로 확인하여야 한다. 확인결과 실제로 행정구역이 변경되었을 때에는 관련 자료를 수정·보완하여야 한다.

이 때 주의하여야 할 점은 그 기본표본조사구에 속한 여러 개의 인구총조사구 중에서 일부의 조사구만 해당되는 경우가 있으며, 극히 드물기는 하지만 1개의 인구총조사구가 2개 동의 경계선으로 나누어 질 수도 있다는 것이다.

나. 新築 아파트 現況資料의 補完

실사표본조사구 내에서 1995년 조사구에서 제외한다. 그 대신에 1995년 11월 1일 이후의 전국의 신축 아파트에 대하여는 별도의 표본조사구를 추출하여 조사하여야 한다.

신축 아파트에 관한 자료는 조사 실시여부와 관계없이 미리 보완하여 놓아야만 조사가 있을 때 표본조사구의 추출기간을 단축시킬 수 있다.

통계청(통계기획국 조사관리과 표본2계)에서는 매년 4월말 및 10월말 기준으로 준공검사가 완료된 아파트에 대하여 아파트명, 동수, 규모(총호수), 소재지, 준공검사일 등의 자료를 각 시·도에서 제출받아 신축 아파트 현황자료를 정리하고 있으며, 늦어도 6월말 및 12월말에는 수집·정리가 완료되고 있다. 통계청에서 정리한 신축 아파트 현황 자료는 담당자에게 의뢰하면 제공받을 수 있다.

신축 아파트 관련 자료는 통계청의 신축 아파트 현황 자료의 내용에 따라 매년 2회 주기적으로 보완·정리하여 두어야 한다. 이에 관한 구체적인 절차와 방법은 'VI. 新築 아파트에 대한 措置'에 자세히 설명되어 있다.

다. 撤去 調査區에 대한 措置

조사과정에서 조사구 내의 주택이 재개발 및 재건축사업, 도로의 신설 등으로 완전히 철거되었을 때에는 조사가 불가능하게 된다. 이 때에는 기본표본조사구

내의 다른 인구총조사구로 교체하여 조사한다. 그리고 기본표본조사구를 추출할 때에 사용한 인구총조사구의 전산 파일에서 그 기본표본조사구의 앞 또는 뒤의 인구총조사구를 기본표본조사구로 추가 포함시키고 처음의 조사구는 추출되지 않은 것으로 취급한다. 그 기본표본조사구 내의 다른 인구총조사구도 모두 철거된 경우에는 그 기본표본조사구의 기본표본조사구를 추출할 때에 사용한 인구총조사구의 전산 파일에서 그 기본표본조사구의 앞 또는 뒤에서 새로 기본표본조사구를 추출하여 기본표본조사구를 교체한다.

라. 非協助 調査區에 대한 措置

극히 드문 경우이기는 하지만 우연하게도 이미 다른 기관에서 조사하여 주민의 협조를 얻기가 어려운 경우에는 기본표본조사구 내의 다른 인구총조사구로 교체하여 조사한다.

2. 留意事項

현지 조사과정에서 기본표본조사구의 유지·관리와 관련하여 발생하는 문제점은 다양하기 때문에 여기서는 과거의 경험에 의하여 기본적으로 중요한 사항만 언급하였다. 따라서 여기에서 언급되지 않은 사항에 관하여는 그때의 사정에 따라 합리적으로 처리하면 된다.

특히 유의하여야 할 사항은 실사표본조사구가 추출되면 조사구요도를 해당 동·읍·면에 송부하여 실제 조사를 할 때에 예견되는 문제점을 수집하여야 한다. 이를 위하여는 실사표본조사구가 추출된 다음 1개월 정도의 기간이 소요되므로, 연구진에서 기본계획을 되도록 빨리 수립하여야만 가능하게 된다.

그리고 기본표본조사구의 추출과 조사구요도의 복사업무는 통계청의 외부 통계정보처리 수입기관인 대한통계협회에 용역으로 의뢰하여 실시하였으며, 이에 관한 업무지침 및 모든 결과물은 관련 문서철(가구단위 통계조사 다목적 표본설계 - 1966)에 보관되어 있으므로 2001년에 2000년도 인구주택총조사의 조사구를 이용하여 기본표본조사구를 개편할 때에는 같은 방법을 따르거나 개선·보완시키면 될 것이다.

□ 著者 略歷 □

● 桂 勳 邦

서울大學校 數學科 卒業

延世大學校 保健大學院 保健學 碩士

現 韓國保健社會研究院 副研究委員

〈主要 著書〉

『1996年 患者調査 改善方案 및 調査設計에 관한 研究』, 1996. (共著)

『OECD의 保健統計 作成基準에 관한 研究』, 1995. (共著)

資料集 96-03

家口單位 統計調査의 多目的 標本設計

1996年 12月 日 印刷

1996年 12月 日 發行

著 者 桂 勳 邦

發行人 延 河 清

發行處 韓國保健社會研究院

서울特別市 恩平區 佛光洞 山42-14

代表電話 : 02) 355-8003

登 錄 1994年 7月 1日 (第8-142號)

印 刷 高麗文化社

© 韓國保健社會研究院 1996
