

人口動態調查 改善을 爲한 方法論的 考察

金 日 炫 · 崔 鳳 鎬

(經濟企劃院 人口統計課)

目 次

I. 序言	IV. 人口動態調查와 關聯된 主要 問題點
II. 人口動態事件을 把握하는 몇가지 調查方法	V. 誤差를 줄이기 위한 方案
III. 經濟企劃院의 人口動態標本調查 事例	VI. 要約 및 結言

I. 緒 言

어떤 集團이든지 그 집단이 갖고 있는 人口規模 (population size or stock)와 構造(structure or characteristics)는 時間의 흐름에 따라 몇가지 基本的인 要因에 의해 不斷히 變化한다. 몇가지 基本的인 要因을 열거하면, 總人口 規模를 變化케 하는 出生·死亡, 婚姻構造 狀態를 變化케 하는 婚姻·離婚, 地域別 人口分析 狀態를 變化케 하는 人口의 移動이 지적될 수 있다. 바로 이렇게 人口의 變動要因이 되는 事件들을 우리는 人口動態事件(vital events)이라 하고, 이를 定理·要約한 것이 人口動態統計(vital statistics)로 定義되는 것이다.

이러한 人口動態統計는 一定時點에서 살펴 보는 總人口規模, 婚姻構造狀態, 地域別 人口分布狀態와 같은 人口靜態統計와 더불어 人口統計의 根幹이 되는 것으로서, 人口政策의 樹立 및 評價, 保健 및 家族計劃事業의 樹立 및 評價, 保險料率 算定에 必要한 生命表의 作成, 어떤 地域에서의 人口動態 要因 把握, 將來 人口推計作業과 기타 여러 研究目的에 緊要하게 活用되는 基本的인 資料인 것이다. 一般的으로 人口動態統計는 動態事件을 직접 經驗한 當事者에 의한 申告를 기초로 作成되는 것이 原則이다. 우리나라의 경우 出生·死亡·婚姻·離婚 申告와 관련하여서는 1909년에 民籍法이, 轉出·入 申告와 관련하여서는 1962년에 住民登錄法

이 만들어 짐으로서 이러한 申告制度가 定着, 現在에 이르기까지 계속 이어지고 있다. 그러나 이러한 申告制度로 부터 完全한 人口動態統計를 作成하는 데는 申告忌避(non-registration), 遲延申告(delayed registration) 또는 內容의 歪曲申告(inaccuracy in registration) 등으로 인하여 制約性이 있다고 할 수 있다.

따라서 申告制度에 의한 人口動態統計를 補完하기 위하여 많은 나라에서 家口를 對象으로 하여 標本調査를 實施하고 있다. 우리나라의 경우 經濟企劃院이 每月 實施하고 있는 人口動態標本調査 및 1987 人口動態 特別調査, 經濟企劃院과 家族計劃研究院(現 韓國人口保健研究院)이 共同으로 實施한 1974 韓國 出生力 調査 및 1983 人口移動特別 調査, 韓國人口保健研究院이 3年 週期로 實施하고 있는 出生力 및 家族保健實態調査와 1988 死亡力 調査 등이 代表的으로 列擧될 수 있다. 물론 이와 같은 標本調査는 申告制度에서 부터 얻을 수 있는 것보다도 人口動態에 관한 精度 높고 深層의(in-depth) 이고 細分된 資料를 提供하여 줄 수 있을 뿐만 아니라, 무단히 變化하고 있는 人口·社會의 特性을 時宜性 있게 把握하게 할 수 있는 重要한 資料出處가 되는 것이다.

그러나 우리가 반드시 理解하여야 할 것은 標本 調査라 하더라도 人口動態事件이 완벽하게 把握될 수는 없고, 나름대로 限界性이 있다는 점이다. 따라서 本稿에서는 經濟企劃院 調査統計局이 每月 實施

하고 있는 人口動態標本調査를 事例로 하여 人口動態事件을 把握하는 調査方法, 調査와 관련된 誤差問題 등을 살펴보고자 한다. 이렇게 살펴 봄으로써 우리는 結果分析 및 類似한 標本調査의 實施를 함에 있어 도움을 얻을 수 있다고 본다.

II. 人口動態事件을 把握하는 몇가지 調査方法

一般的으로 人口動態事件은 몇가지 特徵이 있다. 첫째로 人口動態事件은, 그중에서도 死亡의 경우는 모든 사람이 매년 經驗하는 것이 아니고 極少數만이 經驗하게 되는 發生頻도가 매우 稀少한 事件이다. 예를 들어 粗出生率이 人口 1,000名 當 20名, 嬰兒死亡率이 出生兒 1,000名 當 15라고 假定할 때 100萬名을 調査한다면 1年 동안의 嬰兒死亡을 불과 300件 만을 把握할 수 있을 뿐이다. 둘째로, 人口動態事件은 個人으로 볼 때에는 豫測可能性이 없이 不規則的(unexpected)으로 發生하지만, 한편으로 全體의인 측면에서 볼 때에는 繼續的(continuous)으로 이루어지고 있다는 特徵이 있다. 셋째로는, 人口動態事件의 發生原因이 매우 多樣하다는 점도 지적할 수 있다.

이와같은 特徵때문에 人口動態事件을 어떤 調査를 通하여 把握하는 데는 많은 人力과 豫算의 所要됨과 동시에 事件自體의 捕捉도 어렵다고 볼 수 있다. 특히 死亡의 경우 出生과는 다르게 이미 調査對象이 存在하는 것이 아닐 뿐만 아니라 記憶하기 싫어하는 傾向이 있기 때문에 事件捕捉이 더욱 어렵다. 이러한 特徵을 가진 人口動態事件을 調査하는 몇가지 알려진 方法을 다음에서 살펴 보기로 한다.

1. 1回 回顧調査(Single round retrospective survey)

이는 調査對象家口를 계속 觀察하여 어떤 特性을 把握하는 것이 아니고, 1회에 限해서 調査하는 方法으로서 많은 標本調査가 이 方法에 의해 實施되고 있다. 즉, 이 方法은 어떤 時點에서 調査對象家口를 選定하여 調査對象期間(보통 1年)동안 發生하였던 出生·死亡·婚姻·離婚 및 移動事項을 回顧的

(retrospective)으로 把握하게 하는 方法인 것이다.

그러나 이와 같은 1回 回顧調査에 의한 人口動態事件 把握은 몇가지 要因으로 因하여 完全率이나 正確성에 있어 다음과 같은 誤差가 있는 것으로 알려져 있다.

첫째로는, 時間認識의 差異로 인하여 일어나는 誤差(dating error)로서, 調査對象期間에 發生되어 把握되어야 할 人口動態事件이 調査對象期間에 포함되지 않는 것으로 認識되어 漏落되는 수가 있다.

둘째로는, 어떤 人口動態事件을 經驗한 應答者가 記憶을 못하거나(recall lapse error) 또는 故意的으로 事件을 隱蔽함으로 인해서 事件이 捕捉되지 않을 수가 있다. 특히 出生後 곧바로 死亡한 경우 이것이 醫師나 助産員 등의 失수로 發生하였다면 이들이 이를 故意的으로 隱蔽하기 쉽다. 또한 應答者도 出生後 곧바로 死亡한 아이를 出生兒로 생각하지 않고 死産한 것으로 看做하여 出生·死亡 모두 把握되지 않는 傾向이 높다.

셋째로는, 出生·死亡事件의 發生確率が 극히 작기 때문에 調査員이 面接함에 있어서 每家口마다 심현을 기울여 調査하는 것이 아니고 事件이 없을 것이라고 當然視 여기는 惰性이 생겨 出生·死亡이 제대로 把握되지 못하는 경우가 많다.

넷째로, 死亡의 경우 혼자 單獨家口를 形成하여 살던 사람이 死亡한 경우나 또는 死亡으로 인해서 家族이 解體된 경우는 1回 回顧調査 方式으로는 거의 死亡把握이 어렵다고 볼 수 있다.

2. 多回 回顧調査(Multi-round retrospective survey)

이 方法은 1回 回顧調査에서의 問題點을 補完하기 위한 方便으로 점차 많은 나라에서 이용하는 調査方式이다. 現在 經濟企劃院 調査統計局에서 實施하고 있는 人口動態標本調査(Continuous Demographic Survey)도 이 方法을 適用한 例이다. 즉, 이 方法은 基礎的으로 첫 調査에서 調査對象家口에 대한 여러 人的 事項을 調査하고, 그後 一定期間마다 同 家口를 계속 訪問하여 出生·死亡·婚姻·離婚·轉出·入 등의 變動事項을 記錄

管理함으로써 人口動態事件을 把握하는 方法인 것이다.

따라서 이 方法은 1回 回顧調査에서의 短點 즉, 調査對象期間(reference period)이 길기 때문에 일어나는 記憶誤差나 漏落 등이 減少될 수 있지만 또한 몇가지 問題點이 불가피한 實情이다. 즉, 調査對象家口를 계속 訪問함으로써 應答者의 부담이 크고, 家口(員)의 轉出·入이 심할 경우 調査員의 業務量이 많아지고, 時間이 흐를수록 差別的 人口移動과 經濟社會의 여건의 變動으로 인하여 標本이 老朽化되는 問題點이 內在되어 있다.

이러한 問題點과 관련하여 多回 回顧調査를 實施하고자 할 때에는 몇가지 고려하여야 할 事項이 있다. 첫째로는 調査週期를 얼마로 하여야 하는지 하는 것인데, 보통 3개월 以下가 適正한 것으로 알려져 있다. 둘째로는 調査基準를 常住概念(de jure)으로 할 것이냐 아니면 現在概念(de facto)으로 할 것이냐 하는 問題인데, 보통 常住概念을 채택하고 있다. 셋째로는 어떤 動態率을 計算하기 위해서는 分子가 되는 動態件數와 分母가 되는 人口數가 반드시 必要로 된다. 이와 관련하여 考慮하여야 할 점은 家口의 轉出·入으로 인한 事項과 動態事件이 把握되기 前에 應答者가 먼저 死亡한 事項이다. 즉, 家口의 轉出·入과 관련하여서는 轉出家口에서 情報를 얻기 힘들기 때문에 調査對象을 轉入家口로 代替하게 되는데, 이때 轉入하기 前 發生하였던 動態事件이 漏落되기 쉽기 때문에 이에 대한 細心한 考慮가 要求된다. 또한 出生어 把握되기 前, 母가 死亡한 경우 이를 모두 調査하도록 하여야 한다.

그리고 多回 回顧調査의 變形으로 어떤 標本地域에서 轉出을 하더라도 계속 追跡케 하여 調査하는 方法도 있다. 이러한 方法에 대해서는 用語의 統一이 안되었지만 追跡(follow-up)調査, 縱斷面的(longitudinal)調査 또는 파넬(panel)調査라 알려져 있다.

이 追跡調査 方法의 特徵 가운데 하나는 時間의 흐름에 따라 어떤 個人이나 家口에 있어서 變化되

는 身分이나 構成狀態의 變化分析을 可能케 한다는 點이다. 예를 들면 同 調査方法을 적용함으로써, 어떤 家口가 形成·解體되는 過程, 家族構成員의 經濟活動狀態의 變化, 職業이나 從事上의 地位狀態의 變化에 대한 事項이 쉽게 把握될 수 있다. 그러나 同 方法에서의 문제는 轉出家口에 대한 情報把握이 쉽지 않아 時間이 흐를수록 調査對象家口가 점차 減少하여 標本의 代表性이 떨어지는 限界性이 있다. 따라서 轉出한 家口에 대한 正確한 住所의 把握, 調査員에 대한 충분한 旅費支給 및 應答者가 계속하여 協助할 수 있는 方案 등에 대한 講究가 필요하다.

3. 重複調査(Dual Method Survey)

既存資料(보통 申告資料가 해당됨)가 있을 때, 이 資料와 相互比較 分析하기 위한 目的으로 標本 調査를 實施하는 경우 이를 重複調査 方式¹⁾이라 한다. 즉, 이 方法은 어떤 地域에 있어서 人口動態申告資料 및 轉出·入 申告 資料가 具備되어 있을 때, 다시 同 地域에 대하여 標本 調査를 實施, 兩 結果를 對照 分析함으로써 각각의 申告資料와 標本 調査 結果를 評價하고 漏落된 件數를 把握하게끔 하는 方法이다.

이러한 重複調査 方式은 상당히 단순하게 보인다. 그러나 重複調査에서 좋은 結果를 얻기 위해서는 몇가지 條件이 充足되어야 한다. 즉, 人口動態를 把握하여 相互比較하기 위한 兩 制度는 相互獨立的(independent)이어야 하고, 또한 반드시 같은 地域(same universe)을 對象으로 하여야 한다는 條件이 要求된다. 그러나 現實의 相互獨立的이어야 한다는 條件이 充足되기는 쉽지 않다. 現實의 申告에서 漏落되기 쉬운 動態事件은 마찬가지로 調査에서도 捕捉하기가 어렵다고 볼 수 있다. 이 말은 兩 制度間에 으레히 相關關係가 있기 때문에 相互獨立的이어야 한다는 條件의 違反되기 쉽다는 뜻이다.

예를 들어 살펴보면, 申告上에서 漏落된 嬰兒死亡이나 혼자 살다가 死亡한 경우 또는 집없는 사람

1) 이 方法은 Chandra와 Deming이 처음 試圖한 것으로 찬드라-데밍 方法이라고 한다. 이 方法에서 主된 論點은 어떤 地域에서 發生한 全體動態件數 = 申告나 調査 양쪽에서 모두 把握된 件數 + 申告에서만 把握된 件數 + 調査에서만 把握된 件數 + 양쪽에서 모두 漏落된 件數라는 等式에서 양쪽에서 모두 漏落된 件數를 推定하는 것이다.

의 주민등록 신청을 調査 목적으로 捕捉하는 調査 방법을 할 수 있다. 以外에도 어떤 地域에 具體적으로 集中하지 않지만, 住民登錄이 되어 있어 申告할 場合 등으로 問題가 發生할 수도 있다. 모인 量이 많은 場合 對照分析하는 作業 자체가 많은 時間을 要求하는 여러 問題가 있다.

4. 多目的 標本調査(Multi-subject survey)

앞에서 人口動態事件을 捕捉하는 몇가지 調査方法을 살펴 보았는데, 사실 人口動態統計가 重要하다면 하나라도 이 目的을 위해서 多回 回顧調査制度和 같은 것을 운영하기는 어렵다. 따라서 同一 標本家口를 對象으로 여러 目的을 위한 標本設計을 하게 되는데 바로 이렇게 設計된 標本을 多目的 標本(multi-subject sample)이라 한다.

現在 經濟企劃院 調査統計局의 標本은 人口動態 調査 目的만이 아니고 經濟活動狀態變化, 家計支出(都市地域에서만 調査)의 把握 및 社會指標(social indicators) 作成 등의 目的을 위해 運用되고 있다. 다른 나라의 경우를 살펴보면, 印度에서는 National Sample Survey, 美國에서는 Current Population Survey, 호주에서는 Population Survey, 태국에서는 Integrated Sample Survey 라는 명칭 아래 地方事務所를 運營하면서 多目的 標本調査形態를 취하고 있다.

이와 같은 多目的 標本은 調査費用의 切感, 人力浪費의 防止 등 長點이 많지만 다음대로 問題點이 없는 것은 아니다. 즉, 많은 調査目的을 充足시키기 위한 標本設計上의 問題(design problem), 單一標本을 여러 組織이 같이 使用함으로써 發生할 수 있는 組織間的 摩擦問題(organization problem), 여러 調査를 同一 標本에서 實施함에 있어 調査員에 대한 集中的인 訓練 및 教育의 不足과 應答者의 많은 負擔 등으로 인한 質的 管理 活動上의 問題(quality control problem) 등이 指摘될 수 있다.

Ⅲ. 經濟企劃院의 人口動態標本調査

1. 沿革

同 調査는 戶籍申告에 의한 人口動態統計 및 住民登錄에 의한 人口移動統計를 補充하기 위하여 1963年 부터 經濟企劃院 調査統計局 主管下에 實施되었다. 同 調査의 基本的인 性格은 多回 回顧調査와 多目的 標本調査 方式을 같이 채택하고 있는 것이다. 즉, 調査員은 最初調査에서 標本으로 抽出된 家口에 대한 人的 事項을 把握한 後, 每月마다 家口의 變動事項을 調査함으로써 人口動態統計를 作成케 할 수 있는 方法이다. 만약 調査員에 의한 調査가 아니고 標本으로 뽑힌 家口에 의해서 申告가 된다면, 現在 대만의 家口登錄制度(Household Registration System)와 매우 類似한 것으로 생각할 수 있다.

따라서 同 調査로 부터 어떤 時點에서의 年齡別 人口, 婚姻狀態別 人口, 教育程度別 人口와 같은 靜態統計와 出生·死亡·婚姻·離婚 및 移動에 관한 動態統計 모두가 生産될 수 있는 것이다.

그러나 同 調査는 당초의 期待와는 달리 資料의 信賴도가 낮은 것으로 評價되어 1969年에 調査가 中斷되었다. 그러다가 1972年에 標本을 대폭 改編하여 調査를 再開하였고, 1977年 부터는 調査週期를 과거의 分期別에서 月別로 轉換²⁾하였고, 最近 1988年 5月 부터는 地域別 統計生産이 可能하게끔 標本을 擴大改編 하여 現在에 이르고 있다.

2. 調査體系 및 機能

人口動態標本調査와 관련된 組織間的 關係를 要約하면 그림1과 같다. 各 組織의 機能을 살펴보면, 우선 人口統計課에서는 同 標本調査에 대한 企劃, 調査票의 設計, 集計表案 作成, 結果分析 및 公表 등의 業務를 담당하고 있다. 그리고 調査管理課에서는 標本設計, 地方事務所의 業務監督, 調査의 質的 向上을 위한 活動, 各 實查課의 要求에 대한 調整業務, 標本對象處에 대한 調査協助를 위한 弘報活動, 調査員 訓練에 관한 事項 등을 담당하고 있다.

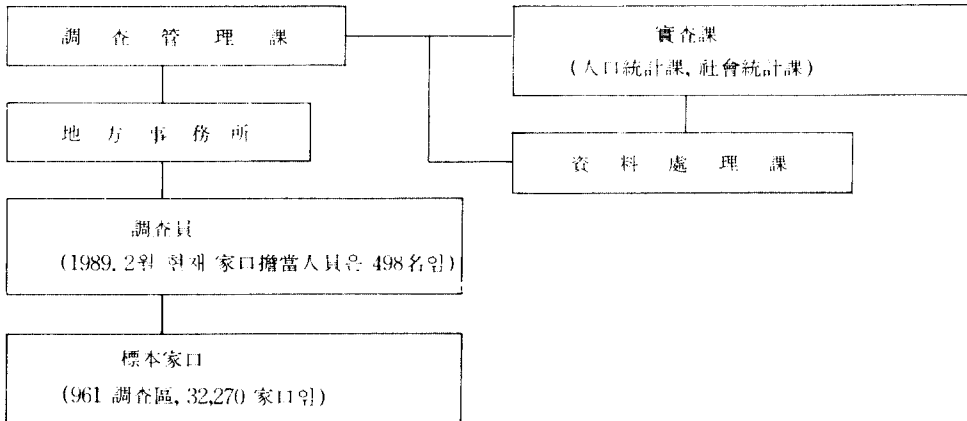
各 地方事務所에서는 調査員에 의해 蒐集된 資料의 內容檢討 및 本部로의 送付, 中央의 調査指針에 따른 細部計劃의 樹立 및 執行, 調査員의 業務指導

2) 또한 1977年 부터 婚姻 및 離婚에 관한 事項과 出生漏落 防止를 위한 妊娠與否狀態를 把握하기 始作하였다.

및 監督 등을 담당하고 있다. 그리고 各 調査員은 每月 標本對象處를 訪問하여 調査業務를 수행하고

있다.

〈그림 1〉 調査體系 흐름圖



3. 標本現況

標本家口를 選定하기 위하여 처음으로 한 作業은 1985年 人口 및 住宅 센서스의 약 151,000 調査區 (市部: 97,000, 郡部: 54,000) 중에서 961個 調査區 (市部: 730, 郡部: 231)를 抽出한 것이다. 그다음 이렇게 抽出된 調査區를 다시 여러개의 集落(1985年 센서스 家口數로 평균 약 10家口를 하나의 集落 基準으로 하였음)으로 分割하여 市部에서는 3個의 集落을, 郡部에서는 4個의 集落을 調査集落으로 選定하고, 選定된 調査集落內에의 居處에 居住하고 있는 모든 家口를 調査對象으로 하였다.

結果적으로 1988年 6月 현재 調査對象家口는 32,209家口(人口數는 118,584名)가 되어 全體의인 平均 抽出率은 1/353이 된다. 그러나 同 標本은 各市·道別 基準으로는 自體加重標本이 되나, 全國의인 資料를 推定할 때에는 各市·道別로 抽出率이 다르게 되어 있어 이를 勘案하여야 한다. 反面 1982年에서 1988年 3月까지 使用되었던 標本은 570個 調査區(市部: 447, 郡部: 93)를 對象으로 하여, 結果적으로 全體抽出率이 1/547인 約 17,000家口의 自體加重標本이 되었다.

4. 調査項目

人口動態標本調査에서의 調査項目은 家口(員)에 관한 事項, 家口(員)의 轉出·入에 관한 事項과

人口動態에 관한 事項으로 區別하여 살펴볼 수 있다. 各 調査項目은 다음과 같다.

① 家口(員)에 관한 事項:

住宅形態, 住宅所有關係, 이름, 家口主와의 關係, 性別, 出生年月日, 學力, 婚姻狀態, 初婚年月, 理想子女數, 妊娠年月, 總出生子女數

② 家口(員)의 轉出·入에 관한 事項:

轉入年月日, 轉入前居住地(區·市·郡 까지만), 轉入事由, 轉出年月日, 轉出地, 轉出事由 등이 基本的으로 調査되고 아울러 轉出·入 家口員에 대한 事項이 把握되기 때문에 移動者와 非移動者間의 特性까지 分析될 수 있다.

③ 人口動態에 관한 事項:

出生에 대해서는 母의 이름, 出生兒의 이름, 出生年月日, 性別, 出生當時 母의 年齡, 出產順位, 妊娠期間, 母의 教育程度, 分娩介助者 등이고,

死亡에 대해서는 死亡者의 이름, 性別, 生年月日, 教育程度, 職業, 妊娠狀態, 死亡原因 및 死亡診斷者 등이다.

婚姻에 대해서는 新郎과 新婦의 이름, 年齡, 教育程度 및 婚姻年月 등이고,

離婚에 대해서는 男便·妻의 이름과 年齡, 婚姻年月, 離婚年月 및 生存子女數 등이 調査項目이 된다.

IV. 主要問題點

統計資料의 質(quality)은 세가지 要素(수, 正確性(accuracy), 合目的性(relevancy) 및 時宜性(timeliness)에 의해 結定되어 지는 것이다. 즉, 어떤 調査를 實施하는 目的이 妥當하여야 되고, 正確하게 蒐集되어야 하며 또한 結果가 빠르게 利用될 수 있도록 公表되어야 하지만 統計의 生命인 質의 條件이 充足되었다고 볼 수 있는 것이다. 이러한 세가지 要素는 어느 한 段階만이 아니라 企劃, 實査, 資料處理, 分析 등 모든 調査段階와 密接한 關係를 맺고 있다고 할 수 있겠으나, 여기에서는 주로 實査段階에서의 正確성과 관련된 側面을 살펴 보고자 한다.

1. 標本問題

現行 人口動態調査를 위한 標本은 1985年 人口 및 住宅 센서스에서 抽出된 調査區를 다시 여러개의 集落(cluster)으로 分割하여 이중 3個(郡部에서는 4個)의 集落內에 居住하고 있는 家戶들이 調査對象이 된 것이다. 그런데 이와 관련하여 問題는 3個의 集落이 相互隣接(contiguous)되어 있어 集落內 家戶들 間의 同質性을 나타내 주는 集落內 相關係數(intra-class correlation coefficient)가 높게 있을 수 있는 점이다. 따라서 現行 標本規模가 32,000家戶나 되는 大規模의 標本이더라도 實際로는 20,000家戶의 역할만 수행할 수 있어 이에 대한 分析이 必要로 된다.

또하나의 問題는 現行 標本은 多目的 標本이기 때문에 각각 다른 目的을 갖고 있는 各 調査가 要求하는 條件을 모두 充足할 수가 없는 短點이 있을 수 있다. 또한 標本抽出時 特別調査區 및 寄宿舍 등의 施設調査區 등이 고려되지 않아 結果分析을 함에 있어 이를 감안하여야 하는 번거로움도 있을 수 있다.

2. 調査完全率上的 問題(Coverage Error)

調査完全率上的 問題는 調査되어야 할 對象이 完全히 調査되지 못하고 漏落(omission)되거나 重複(duplication)되게 調査되는 現象을 意味한다. 여기에서 調査되어야 할 對象은 調査性格에 따라 人口 및 住宅 센서스의 경우 家戶(員)·家戶·住宅 등이 해당되고, 人口動態調査의 경우 人口動態事件(出生·死亡·婚姻·離婚·轉出·入)이, 經濟活動調査에서는 經濟活動人口가, 그리고 障者調査인 경우는 장애인 對象으로 分類될 수 있다.

대체로 調査完全率上 問題가 發生되는 原因으로는 調査員의 잘못 以外에 누가 調査對象이 되고 또 어떤 方法으로 把握하여야 한다는 基準이 不明確한 때나, 調査準備段階에서 作成된 調査區 要圖(enumeration area map)나 家戶名簿(household list)가 不完全한 경우에 發生할 수 있다.

따라서 調査가 完了되면 이와같은 調査完全率上의 問題를 檢討함이 一般의이다. 이를 把握하는 方法으로는 몇가지가 있는데 첫째로 再調査³⁾를 하는 것이고, 둘째로는 過去調査資料와 結果를 比較하는 것이고, 셋째로는 人口學의 技法을 活用하여 分析하는 方法,⁴⁾ 넷째로는 住民登錄 및 戶籍申告 등의 行政報告 資料와 比較分析하는 方法 등을 用기할 수 있다.

3. 無應答 誤差 問題(Non-response Error)

이 問題는 調査對象處에 대한 情報獲得이 部分的으로나 또는 全體적으로 失敗된 경우를 意味하는 것이다. 이의 原因은 몇가지로 그 類型이 나누어진다. 즉, 調査員이 標本家戶를 찾아내는데 失敗하거나(non-contact), 調査員의 訪問時 應答者가 不在中이거나(not-at-home), 應答者가 應答을 忌避하거나(refusal), 完成된 調査票를 紛失하거나(loss) 또는 調査直前에 家戶(員)가 轉出함으로써 情報獲得이 안된 경우도 이에 포함될 수 있다.

따라서 Non-response 問題를 다루기 위해서는

3) 再調査方法은 먼저 投入된 調査員보다 有能한 調査員이 2차의 調査에서 얻어진 結果를 活用하여 評價하는 從屬 調査方法(dependent method)과 먼저 調査와 똑같은 條件으로 再調査를 實施하여 評價하는 獨立(independent) 調査方法으로 區分할 수 있다.

4) 例를 들어 人口의 경우 年齡別로 生殘率을 計算하여 補으로써 人口數의 漏落率을 把握할 수 있다.

먼저 情報獲得이 안된 要因을 把握하는 것이 매우 重要하다. 그런다음 이에 대한 對處를 하는 것이 必要로 되는데 몇가지를 언급하면 다음과 같다.

첫째로는 調査對象가구를 代替(substitution)하는 方法을 생각하여 볼 수 있다. 둘째로는 資料處理 過程중 無應答(Non-response) 誤差가 있는 狀態에서 集計하여 그 結果를 信賴할 만한 推定値와 比較하여 맞지 않는 경우 이를 어떤 加重値(weight)를 주어 調整(adjustment)하는 것으로 많이 利用되고 있는 方法이다. 세번째로는 資料處理 過程중 情報獲得이 안된 部分에 대해서 利用 可能한 他資料로 代替(imputation)하는 것으로 이에 는 콜덱(cold-deck) 方法과 핫덱(hot-deck) 方法이 있다. 콜덱 方法은 現調査에서 漏落된 資料를 同一 形態의 過去調査로 부터 同質的인 特性을 가진 資料로 代替(imputation)하는 方法을 뜻하고, 핫덱 方法은 現調査에서 蒐集된 資料를 利用하여 漏落된 資料를 代替하는 方法을 意味한다.

參考로 無應答(Non-response)에 대한 다른나라의 例를 살펴보면 表1과 같다. 먼저 同表에서 無應答率을 보면 모든 나라가 나름대로 높은 無應答率을 갖고 있는 것으로 나타난다. 한가지 特記할 만한 事項은 先進國에서는 無應答의 非協助(refusal)로

인한 無應答率이 높은 反面, 필리핀이나 터어키와 헝가리 같은 나라에서는 無應答의 非協助는 굉장히 낮은 것으로 나타나는 점이다. 즉, 이들 나라에서는 일단 應答者를 만나기만 하면 情報獲得이 可能한 것으로 볼 수 있지만, 先進國에서는 應答者의 非協助가 보다 큰 問題임을 알 수 있다.

또 다른 無應答에 대한 實例를 美國의 1980年 出産力調査에서 살펴볼 수 있다(表2 參照). 同調査는 既히 接受된 出生申告書(birth certificate)의 內容을 보다 深層의으로 把握하고자 調査票를 새로 作成하여 郵便으로 出生兒의 母와 出生된 病院, 그리고 分娩介助者에게 각각 送付하는 方法을 적용한 것이다. 同表에서 우선적으로 보이는 것은 1次 郵便送付 結果 回收된 率は 病院으로 부터 가장 높았고(50.7%), 出生兒의 母로 부터 가장 낮게(36.0%) 나타났다. 그러나 2次, 3次로 계속 試圖하여 調査를 終結할 때를 살펴보면 出生兒의 母로 부터의 回收率이 높은 反面, 分娩介助者로 부터의 回收率이 가장 낮음을 알 수 있다. 그리고 同表에는 언급이 안되었지만 嬰兒死亡의 경우 無應答率이 出生調査의 경우보다 높아, 調査가 복잡하거나 어려울수록 無應答率이 높아진다는 事實도 지적하고 싶다.

表 1. 各國 調査別 無應答率

(Unit : %)

	Title	Year	Type		
			Total	Non-contact	Refusal
U. K.	General HH Survey	1974	14.1	2.4	11.7
Sweden	Labor Force Survey	1978	6.5	2.9	3.6
	HH Income Survey	1978	12.3	4.2	8.1
Norway	Family Expenditure Survey	1977	31.0	-	-
	Labor Force Survey	1977	14.1	7.5	6.6
U. S. A	Current Population Survey	1977	4.1	1.6	2.5
Philippines	Fertility Survey	1978	14.6	14.2	0.4
Turkey	Fertility Survey	1978	20.8	19.8	1.0
Hungary	Survey on Basic Social-Demographic Data	1988	14.3	10.4	3.9
	Income Survey	1988	16.9	11.7	5.2

Source : UN, National Household Survey Capability Program : Non-sampling Errors in Household Surveys, 1982

表 2. 美國의 1980年 出生力 調査에서의 應答形態

調査對象	標本規模	計	1次 郵便送付	2次 郵便送付	3次 郵便送付	電話	無答
出生兒의 母	7,825	100.0	36.0	17.1	3.2	23.2	20.5
病院	9,855	100.0	50.7	18.3	-	7.1	23.9
分娩介助者	7,939	100.0	41.1	14.5	-	6.0	38.4

Source : National Center for Health Statistics, Methods and Response Characteristics : 1980 National Natality and Fetal Mortality Survey, Series 2, No. 100

4. 應答誤差((Response Error) 問題

應答誤差라 함은 얻어진 情報가 不正確한 경우를 意味한다. 즉, 應答誤差는 應答者의 잘못뿐만 아니라 調査票設計, 調査指針書 作成, 調査員 訓練 등을 담당 企劃하는 사람 모두에게 달려 있는 것이다. 이의 原因을 보다 具體的으로 살펴보면 不適切한 調査票의 設計, 不適切한 調査員 訓練, 調査員의 調査에 臨하는 態度, 應答者가 調査에 臨하는 態度, 適切한 應答者가 아닌 사람에 對해서 적당히 情報를 얻어내는 경우(proxy response) 등이 指摘될 수 있다.

應答誤差는 다시 어떤 無作爲的으로 發生되는 (resulting from transient or random factors) response variance 要因과 調査與件에 起因되는 (due to the essential survey conditions that are inherent in the design of the survey and the general setting) response bias 要因으로 나누어 질 수 있다.

response variance는 만약 똑같은 標本과 節次下에 標本調査를 두번 實施한다고 하더라도 兩 調査의 推定値는 差異가 있을 수 있는데, 바로 이 差異가 response variance 要因 때문이다. 쉬운 用語로 이 要因은 無意識的인 것으로 分類할 수 있다.

反面 response bias는 다소 意識的인 것으로 이것이 發生하는 상황을 살펴보기로 한다. 만약 두個의 調査를 實施한다고 한 때, 한 調査에서는 應答者의 電話番號를 記入하게 하고 署名捺印을 하게 하지만 또는 調査結果에 영향을 미칠 수 있는 第三者가 지켜보고 있는 경우와, 다른 調査에서는 이렇지 않은 경우를 比較하여 볼 때 後者의 結果가 보다

참값(true value)에 接近할 수 있다. 여기서 兩 調査 結果間의 差異를 바로 response bias 때문으로 볼 수 있는 것이다. Response bias의 例로는 年齡의 잘못報告, 出生子女數의 漏落報告(under-reporting), 人口動態事件의 漏落報告(omissions), 記憶誤差(recall lapse)로 인하여 調査對象期間을 混同하여 잘못 調査되는 경우, 追跡調査時 應答者의 喪증이 誘發되어 發生하는 誤差(panel bias), 不適切한 應答者로 부터 情報를 얻는 過程에서 잘못 調査되는 경우 등을 열거할 수 있다. 이 가운데 不適切한 應答者로 부터 情報를 얻는(proxy response) 比率를 다른나라의 例로부터 살펴보면 表 3과 같다. 이 表에서 우리는 우리나라가 아닌 이태리의 예이지만 家口調査時 調査目的에 맞는 適切한 應答者가 아닌 老人이나 어린이 또는 옆집에 사는 사람같은 不適切한 應答者를 통해 調査를 하고 있는 比率이 상당히 높음을 알 수 있다.

表 3. 不適切한 應答者(proxy response)의 比率 : Italy

Title	Year	%
Family Behaviour Survey	1984	52.2
Health Survey	1983	37.5
Work Force Survey	1985.7	38.9
Work Force Survey	1985.10	39.0

Source : Vittorio Buratta, The Problem of Proxy Responses in Household Surveys, Paper Presented at the First Conference of International Association for Official Statistics held in Rome, 1988

V. 誤差를 줄이기 위한 方案

以上에서 살펴 본 誤差는 統計調査를 實施함에 있어 없어야 하는 것이 理想的이 되겠지만, 實際 얼마간의 誤差가 存在하게 되는 것은 불가피한 實情이다. 그렇다고 하더라도 우리는 誤差를 줄이기 위한 努力을 계속해 해서는 안된다. 그리고 誤差는 誤差發生의 事前 防止活動과 分析作業이 가장 重要한 것이다. 이와 관련하여 다음과 같은 네가지를 생각하여 보았다.

1. 一般的인 品質管理活動 (Formal quality control)

이는 誤差發生의 事前 豫防의인 측면으로, 잘 알려진 바와 같이 ① 철저한 用語 및 分類의 使用, ② 事前調査 및 試驗調査의 實施, ③ 調査票 設計의 철저, ④ 調査指針書 作成의 철저, ⑤ 標本틀(frame)의 正確한 把握, ⑥ 適切한 調査對象期間 및 調査基準日 設定 등을 통한 徹底한 調査企劃과 遂行이 당연히 포함된다.

아울러 教育訓練의 強化, 調査員 監督 및 指導體系의 強化, 効率的인 實查指導 實施, 調査員의 士氣 振作, 調査環境의 積極적 改善 등도 資料蒐集과 관련된 品質管理活動이라 할 수 있다.

2. 調査方法의 檢討

調査方法은 調査員이 質問하여 記錄하는 面接他記式과 郵便이나 調査員에 의해 배부된 調査票를 應答자가 직접 記入 作成케 한후 郵便이나 調査員이 회수하는 自記式으로 나누어질 수 있다.

面接他記式의 長點은 應答자의 높은 協助와 應答率을 얻을 수 있고 또한 보다 完壁하고 一貫性있는 資料를 얻을 수 있는 利點이 있지만, 이 方法은 일반적으로 많은 費用과 人力을 必要로 하고 또한 調査員에 대한 많은 訓練을 必要로 하는 短點이 있다.

한편 自記式은 적은 費用과 人力으로 調査를 實施할 수 있고, 많은 家族構成員이 相互 協議하여 調査票를 作成할 수 있는 機會를 주게 되어 한 사람이 全體 構成員에 대한 應答을 하는 것보다 內容이 더 正確할 수 있는 長點이 있을 수 있다. 短點으로는 調査票의 回收率이 낮아 電話나 再訪問이 必要

로 되는 것과 資料處理 단계에서 內容檢討量이 많다는 것이 지적될 수 있다.

우리나라에서는 現在까지 調査의 대부분이 面接他記式에만 의존하여 온 경향이 있다. 그러나 점차 家口調査時 忙시간에 適切한 應答者를 面接하기가 어려운 것으로 豫想된다. 따라서 이 時點에서 面接他記式만을 채택한 것이 아니고 兩 調査方法의 長·短點을 살펴 보아 轉換도 생각할 必要가 있다. 日本이나 호주에서 직접 調査員이 調査票를 配付하고 다시 직접 回收하는 方法도 肯定的으로 研究할 必要가 있다.

3. 事例集의 發刊

이는 調査를 企劃하고 資料를 蒐集하고 資料處理를 하는 過程에서 發見된 問題는 무엇이고 또 이를 解決한 方法은 무엇이었다 하는 事例들을 모아 整理할 必要가 있다는 것을 뜻한다. 즉, 無應答은 얼마인데 이를 어떻게 處理하였나 하는 事項, 內容檢討中에 發見된 問題點 및 處理結果, 資料處理 過程中에 發見된 問題點 및 解決方案 등이 整理되고 또한 다른 사람에게도 알려질 必要가 있다.

4. 標本設計에 對한 分析

이런 母集團에 대하여 總數나 平均 등의 統計量을 각각 推定할 수 있도록 標本을 여러개의 副次標本(sub-sample)으로 나누어 調査를 試圖하여 結果를 分析하는 것, 最終 標本抽出單位(ultimate sampling unit)의 規模에 대한 研究, 標本이 과인 意圖하였던 대로 되었다 하는 檢證이 標本設計과 관련시켜서 생각할 必要가 있다고 본다.

VI. 要約 및 結言

以上에서 限定된 知識과 經驗을 가지고 人口動態 調査를 中心으로 人口動態統計의 重要性, 調査의 種類, 事例, 誤差의 種類 및 誤差를 줄이고자 하는 努力들에 대하여 살펴 보았다.

간단히 要約하여 보면 人口動態統計는 人口·保健政策의 樹立 및 評價, 生命表의 作成, 將來人口推計作業 등에 緊要하게 活用되는 資料인 것이다. 이러한 人口動態統計는 申告에 의해서 作成됨이 원칙적이겠으나, 申告制度에는 申告忌避, 遲延申告

또는 內容의 歪曲申告 등으로 인하여 制約性이 內在되어 있기 때문에 이를 補完하고 보다 精度 높고 深層的인 資料를 얻기 위해서 人口動態에 관한 標本調査를 實施하는 경우가 많다.

調査의 方法은 人口動態事件의 몇가지 特徵 때문에 다른 調査와는 달리 調査方法이 몇가지로 나누어질 수 있다. 즉, 1회 回顧調査(single round retrospective survey), 多回 回顧調査(multi-round retrospective survey), 重複調査(dual method survey) 方法으로 細分되어지는 것이다. 現在 經濟企劃院 調査統計局이 每月 實施하고 있는 人口動態 標本調査는 바로 多回 回顧調査 方式을 채택하고 있는 것이다.

그러나 이러한 方法으로 調査를 하더라도 人口動態事件의 件數와 內容이 완벽하게 把握될 수는 없고 나름대로 限界性이 있음을 理解하여야 한다. 이와 관련된 問題點을 標本問題, 調査完全率上의 誤差(coverage error), 無應答의 問題(non-response error) 및 應答誤差(response error) 類型으로 區分하여 살펴 보았다. 여기서 한가지 강조할 事項은 이러한 誤差는 없어야 하는 것이 理想的이겠지만, 現實적으로 얼마간의 誤差는 存在한다는 것이다.

그럼에도 불구하고 우리는 誤差를 極小化하기 위한 努力을 게을리 해서는 안된다는 點이다. 이러한 努力의 일환으로 앞에서 一般的인 品質管理活動(formal quality control)의 強化, 調査方法의 再檢討, 事例集의 發刊 및 標本設計에 대한 分析 側面을 다루어 보았다.

끝으로 茶山 丁若鏞(1762~1836)이 牧民心書에서 強調한 事項 몇가지를 言及하면서 마무리 짓고자 한다. 그가 살던 時代에서 까지도 가장 重要한 일종의 하다가 良賤身分의 混肴防止 및 賦稅·搖役을 지우기 위한 戶籍簿의 作成이었다. 그러나 이 당시의 戶籍制度는 극히 不完全하여 이로 인한 百姓의 피해는 극심한 것이었다. 茶山은 이러한 不完全한 制度의 改善을 건의하였는데, 이것이 바로 統計調査의 實施를 의미하는 것이었다. 즉, 그는 한 고을의 首領이 된 者는 實際 그 고을에

살고 있는 家口의 特性과 生活狀態를 詳實하게 調査하고 이를 기초로 經緯表를 作成하여 善政에 用될때 活用되도록 提示하였다. 이런 過程中에 그는 이미 오늘날과 같은 地圖作製의 必要性과 非標本誤差를 줄이기 위한 方案 등을 提議하였고, 끝으로 調査를 담당하는 者의 誠實性을 가장 強調한 바 오늘날에도 示唆하는 바가 적지 않다.

參考文獻

- 經濟企劃院, 調査統計局, 非標本誤差의 最小化 方案, 1988
- 經濟企劃院 調査統計局, 人口動態標本調査 指針書, 1988
- 經濟企劃院 調査統計局, 家口對象調査 家口管理 指針書, 1988
- 金景中, 韓國의 經濟指標, 每日經濟新聞社, 1989
- 民族文化推進委, 國譯 牧民心書, 1969
- Buratta, Vittoria, The Problem of Proxy Responses in Household Surveys, paper presented at 1st Conference of International Association for Official Statistics held in Rome, 1988
- ESCAP, Review of Demographic Sample Surveys in Asia and the Pacific:1970-79, 1980
- Kiregyera, Ben, Quality Problems Connected with Field Data Collection, paper presented at 1st Conference of International Association for Official Statistics held in Rome, 1988
- National Center for Health Statistics, Methods and Response Characteristics:1980 National Natality and Fetal Mortality Surveys, Series 2, No. 100, 1986
- Purakam, N., Problems in Designing an Integrated Sample Survey System in Thailand, paper presented at the 42nd Session of International Statistical Institute held in Manila, 1979

〈Abstract〉

A Study on Vital Statistics Survey : its Type, Source of Errors and Improvement Scheme

Il-Hyun Kim · Bong-Ho Choi

(National Bureau of Statistics Economic Planning Board)

It is well known that vital statistics is of great importance as basic data for establishing various range of national policies. Especially, vital statistics is important among demographic information for monitoring and evaluating the population policy, for constructing life table, for making population projection, and for studying various aspects of the society.

In principle, the production of vital statistics is based on the registration system. It is, however, still observed that there are some limitations in utilizing fully the registration system due to the inherent problems such as problems in its coverage, accuracies and timeliness.

Thus, as an alternative, many countries conduct survey on vital statistics in order to supplement the registration system and obtain in-depth data. Korea is no exception in this aspect. The National Bureau of Statistics carries out the so-called Continuous Demographic Survey. This is a kind of multi-round retrospective survey, covering 32,000 households and having reference period of one month. The survey has also characteristics of multi-subject sample. Thus, surveys on economic activity status of population, house-hold income & expenditure, and social indicators are together conducted with the same sample.

It is, however, found that the survey itself tends to have some quality problems. Especially, the quality problems connected with field data collection are summarized as coverage error, non-response error and response error.

Although it is inevitable not to be free from these errors, we should make all our efforts to reduce the errors. The probable schemes pointed out in this paper are as follows : 1) the strengthening formal quality control activities, 2) the review of the survey method, i.e., the combining interview method with mail-sending and mail-back method or pick-up method, 3) well documentation for various cases found in every stage of data collection, and 4) the strengthening the analytical activities. It is, also, emphasized that sincerity of planners and interviewers is the most important factor among other things.