Master Sample について

三 谷 与 司 夫*
小 田 正 也

1. Master Sample の意義

Master Sample は普通「親標本」と訳されている.

調査機関が行う標本調査の Sompling は，多くの場合多段抽出法がとられる．又同様の Sampling手続をとる調査も多い．この様な調査にその都度同様の手続を繰返すのは時間，労力等の無駄な経費を消費する．そこで或段階までの Sampling を行い，相当多量の標本を手元に常備し，必要に応じてこれから更に Sampling する．この或段階までの標本を Master Sample と云う.

2. 消費者調査の Sampling

消費者調査には調査単位が世帯と個人の二種類ある．この様な母集団の枠の作成は殆んど不可能である．そこで官庁資料を利用する．これらの資料としては住民登録，選挙人名簿，国勢調査資料等があるが，いづれも不完全か或は利用困難である.

3. 弊社調査部の消費者標本調査用 Master Sample 作成の手続

当調査部の Master Somple の抽出方法は「層化集落抽出」である．一般の調査はこれから更に Sampling を行うから「層化二段抽出」となる.

設定地域は，大阪市，堺市，布施市，守口市，吹田市，豊中市，神戸市，芦屋市，西宮市，尼崎市，京都市の11都市で調査単位は世帯及び個人である.

層化の基準は原則的に地理的層化に従い，生活環況も考慮に入れた．地域内を一層約15,000世帯均一になる様に町名単位に93に層化した．更に各層を50世帯の集落，300集落に分け，三集落を抽出して Master Sample とした.

実際の手続は国勢調査の町名別世帯数表により，地図の上で網の目状に層化して，「層別町名一覧表」を作成する．町名の配列は地図の上の一定の方向をとる様に列記し，町別世帯数の累積和を算出列記する．次に系統的に3ケの乱数をとり，累積和の数に当つた町を，抽出集落を含む町とする．更に町内の乱数をとり，国勢調査の準備調査で設定された集落（調査区）で，この乱数に相当する集落を抽出集落とする．この集落に属する全世帯を listing して50世帯の集落にまとめる．この時世帯構成員の氏名，年令，職業，性別を listing し個人調査にも利用出来るようにする．これらは I.B.M カードに設計した Master Sample カードに記入保存する.

4. Master Sample への標本割当

標本割当に当つて最も重要な事は集計操作上煩鎖にならない様注意を要する．もしこれを誤ると集計不能にさえなる事がある.

今調査単位の或計量を x とする．母集団の総量の不偏推定値X'は

$$X'=\sum_{i=1}^{L}\left\{\frac{M_i}{m_i}\sum_{j}^{m_i}\left(\frac{N_{ij}}{n_{ij}}\sum_{k}^{n_{ijk}}x_{ijk}\right)\right\}\cdots\cdots\cdots\cdots(1)$$

i: $(1\cdots\cdots L)$ 層番号
j: $(1\cdots M_i)$ 集落番号(当部の場合$M_i=\overline{M}=300$)
k: $(1\cdots N_{ij})$ 調査単位の番号（当部の世帯調査の場合$N_{ij}=\overline{N}=50$)
m_i: 抽出集落個数(当部の場合$m_i=\overline{m}=3$)
n_{ij}: 集落からの抽出標本単位数
$\sum_{j}^{mi}, \sum_{n_{ij}}^{nijk}$ 等は抽出されたものについて加える事を示す

この式において各層からの集落の抽出率

$$f_i=\frac{m_i}{M_i}=\frac{\overline{m}}{\overline{M}}=f=\frac{3}{300}=\frac{1}{100}$$ と一定となる.

$\frac{n_{ij}}{N_{ij}}=g_{ij}$ は集落からの抽出率である．ここで $g_{ij}=g$ と一定とすると(1)式は次の様になる.

$$X'=\frac{1}{f}\frac{1}{g}\sum_{i=1}^{L}\sum_{mi}^{Mi}\sum_{k}^{nijk}x_{ijk}$$

この式は抽出標本を層，集落に関係なく合計出来る事を示している．即ち集計にはこれらを考慮する事なく単純に加算し得る．故に世帯調査の場合には $N_{ij}=\overline{N}=50$ 一定であるから各集落に均等に標本を割当てればよい．個人調査の場合にはN_{ij} は一定とはみなせない．故に n_{ij} は N_{ij} に比例して割当てればよい．いづれにせよ一定の抽出間隔で系統的に割当てればよい．即ち Sampling 手続は非常に簡単である.

5. 調査結果の精度計算

精度は勿論 Sample Size 母集団の分布状況によつて異るが，層化二段抽出の線形推定の場合（即ち $\frac{n_{ij}}{N_{ij}}=g_i$ である場合．この場合はこれに相当する。)

* 株式会社電通大阪支社

の総量の分散 $V\ (X')$ の不偏推定は

$$V(X')=\sum_{i=1}^{L}\left[\frac{M_i(M_i-m_i)}{m_i}\left(\frac{M_i}{M_i-1}\sigma_{ie}^2\right)+\frac{M_i^2}{m_i}\frac{1-g_i}{g_i}\sigma_{iw}^2\right]\cdots\cdots\cdots\cdots(2)$$

$$\sigma_{ie}^2=\frac{1}{M_i}\sum_{j=1}^{M_i}(X_{ij}-\overline{X}_i)^2,$$

$$\sigma_{iw}^2=\frac{1}{N_i}\sum_{j=1}^{M_i}\sum_{k=1}^{N_{ij}}(x_{ijk}-x_{ij})^2$$

ここで σ_{ie}^2, σ_{iw}^2 は標本値から推定しなければならない．標本値の各分散を S^2 で表せば

$$S_{iw}^2=\frac{1}{\mathrm{m}}\sum_{j}^{m_i}N_{ij},\ S_{ij}^2\ ,\quad S_{ij}^2=\frac{1}{n_{ij}-1}\sum_{k}^{n_{ij}}(x_{ijk}-\overline{x}_{ij})^2$$

$$\overline{x}_{ij}=\frac{1}{n_{ij}}\sum_{k}^{n_{ij}}x_{ijk}$$

$$S_{ie}^2=\frac{1}{m-1}\sum_{j}^{m_i}(x_{ij}''-\overline{X}_i'')^2,\quad X_{ij}''=\frac{N_{ij}}{n_{ij}}\sum_{k}^{N_{ij}}x_{ijk}$$

$$\overline{X}_i''=\frac{1}{m_i}\sum_{j}^{m_i}X_{ij}''$$

とする．一方

$$ES_{iw}^2=\sigma_{iw}^2$$

$$ES_{ie}^2=\frac{M_i}{M_i-1}\sigma_{ie}^2+\frac{1-g_i}{g_i}\sigma_{iw}^2$$

従つて

$$\frac{M_i}{M_i-1}\sigma_{ie}^2\approx S_{ie}^2-\frac{1-g_i}{g_i}S_{iw}^2$$

$$\sigma_{iw}^2\approx S_{iw}^2$$

これらを(2)に代入すれば $V(X')$ を得る．個人調査はこれによればよい．

（文献目録52頁より続く）

73. Controllership Foundation, Inc.
"Management Planning and Control, An Annotated Bibliogaaphy" (Seriss II. Business Planning and Control. Report NO. 4) March 1955.

74. Symonds, G.H.
"Linear Programming: The Solution of efineryR Problems," Esso Standard Oil Company, New York, N.Y. 1955

75. Dunlap J.W.
"Would Operations Research Help in Your Company?" N.A.C.A. Bulletin October, 1954

76. Zajac E.C.
"A Case In Control ofMaintenance Labor Costs" N.A, C.A. Bulletin July 1955

77. Magee J.F.
"Operations Research and the Accountant," N.A.C.A. Bulletin, August 1955