

# 統計調査法

関 彌 三 郎

- 一 悉皆調査と標本調査
- 二 統計調査法（統計的方法）
- 三 悉皆統計調査の正確性
- 四 其他の統計調査法
  - (イ) 推算
    - (甲) 標本調査法（一部調査法）
      - ①有意選択法
      - ②任意抽出法（抽出調査法）

## 一、悉皆調査と標本調査

社会的集団現象に就ての統計調査は、徴兵、徴税等の行政上の目的から既に古代より行われて来たのであるが、科学的な統計的方法が発達しそれによつて調査がなされるようになったのは近世以降の事柄である。

近代的な中央集権國家の成立、發展は、その政策樹立の基礎として一國の人口、經濟其他の社會狀態の正確な數量的把握の必要性を増大せしめ、殊に民主政治の採用による國民の代表者たる議員の數の各地方への割當の必要に促されて、十八世紀の末より國勢調査 Census として人口調査が漸次各國に於て行われ、それは後には人口のみならず社會狀態、産業狀態の調査にも及ぼされるようになったのである。

一方社會現象の集團的觀察による一般の性質、規律性の研究即ち社會現象の數量的、統計的研究は、十七世紀の末葉「政治算術 Political Arithmetic」の名の下にイギリスに於て起り、その發展は又社會的集團現

象の統計調査法の進歩を助長し、十九世紀の中頃より「統計学 Statistics, statistics」の名を与えられ、爾來社会現象の研究並に統計的方法の研究を行うようになった。

大體社会的集團現象の正確な認識は、国勢調査のよりにその構成要素のすべてを洩れなく調査し、その大いさ並に構成状態を数量的に明らかにする事によつて可能となるのであるが、それは十八世紀の終りになつて始めて実行に移されるようになったのであつて、それ迄は他の目的でなされた記録（例えば出生、死亡、婚姻、移住等）に関する寺院の記録や行政官庁への届出等）よりする算定や、既知の統計値の加工、誘導による推算、又は一部分のみの調査による全体の推計（標本調査）等の、正確性を欠くが簡便な方法で統計調査はなされていたのである。

しかし社会的集團現象の悉皆調査は、統計調査機関の発達を前提に、多大の費用と労力を費してのみ可能であるためその実行は容易でなく、これ等の犠牲との

關聯に於ては不可能と看做し得る現象が極めて多く、又中にはその性質上全く悉皆調査を許さないものもあるため、その後も推算や標本調査は悉皆調査と並んで社会的集團現象の有力な調査方法であつた。先に述べたように悉皆統計調査は最初人口調査として発達し、

十九世紀の中葉以降次第に經濟方面にも及ぶようになつたのであるが、それは政治活動の必要によるものとはいへ、更に又悉皆調査の手續を人口現象は比較的容易に満すものであるからである。蓋し、人口現象を構成する要素は個々の個人であつてその存在を把握する事は容易であり、且それに就て觀察すべき諸性質は自然的なものが多く明瞭に認知し得るために、それに対する悉皆觀察は左程困難ではないからである。しかるに經濟現象を構成する要素は各個人の日々の生活に於て反覆生起しその数は極めて多く、又瞬間的現象が多いためその把握は容易でなく、且個々の要素に就て觀察すべき性質は社会的なものなるため多義的であり、不明瞭な場合が多いのでその悉皆觀察は非常に

困難である。故に経済現象に於てはそれの統計的認識の政治的、学問的必要性は、他の如何なる社会現象にもまして重大であるにも拘らず、悉皆統計調査の困難な乃至は行い得ない部面が相当大であり、標本調査、推算の役割は極めて重要である。

かくて主として社会調査法としての統計的方法を取扱つて来たドイツ統計学派に於ては、統計的方法は社会的集団現象(「社会大量 soziale Masse」という)の悉皆観察の方法(これを「大量観察法 (Massenbeobachtungsmethode)」といふ)であつて、標本調査や推算是「統計的方法の代用法 Surrogate der Statistik」とされて来たのである。

しかるに一方社会現象、一般的には非類型的現象の集团的、数量的研究方法としての統計的方法を取扱つて来た英米の統計学に於ては、その方法的基礎たる大数法則 Law of large numbers を確率論によつて証明し、それより進んで一部分の現象の集团的観察(従つて標本観察)による、現象の一般的性質、規律性の

認識の誤差の限界を求める事を可能ならしめるに至つたのである。しかしながらその標本論は多くの現象の観察を必要とする「大標本論 Large sample theory」であつたために極めて不便であり、この欠点を除かんとして今世紀の初頭より發達して来たゴセット (W. S. Gosset 筆名 Student) フィッシャー (R.A. Fisher) に始る近代数理統計学は、現象の一般的性質の統計的研究は本質的に部分観察、標本観察の方法であるとして再検討し、観察資料の簡約を如何にして可能ならしめ得るかを考察する事によつて、従来の大標本論を内に含む「小標本論 small sample theory」——正確には「精密標本理論 exact sampling theory」を完成し發展せしめたのである。従来の統計理論が観察資料の整理、分析なる統計的方法の記述的な側面を主として取扱つていたのに対して、この精密標本理論は統計的方法の推測的な側面を重視する点に於て全くその趣を異にし、従つて従来の「記述統計学 descriptive statistics」に對して「推測統計学 inductive statistics」又は従来の一統

「推計学 statistics」に代る「推計学 stochastics」といわれているのである。

そしてこの成果がその生誕地たる農業実験の分野より漸次他の自然科学方面に適用され、第二次世界大戦の直前アメリカ合衆国に於て社会的集団現象の統計調査にも利用されるようになり、社会現象に対する標本調査を正確ならしめるに至つたのである。従来の社会的集団現象に対する標本調査に於ては、その結果が悉皆調査による時得られるであろう結果に対して如何なる關係に於てあるかが不明であつたのであるが、この新しい方法に於てはそれが標本誤差として規定され得るため、標本調査は悉皆調査に劣らず社会的集団現象の正確な認識を可能ならしめるのである。

そしてこの方法は標本構成要素の任意抽出 random sampling を根本条件とする方法であるため「任意標本調査 random sample survey」又は「抽出調査 sampling survey」といわれ、その正確性と簡便、低廉なる事の故に社会統計調査のあらゆる方面に於て急

速に發達し各国に普及して行つたのである。我が国に於ても今次大戦後 G・H・Q の統計家達によつて始めて任意標本調査の技術が導入され、昭和二十一年七月より始る消費者価格調査 Consumers price survey を皮切りに短時日の間に發展を遂げたのである。

しかしながら任意標本調査法の実行には種々の条件が具つていなければならないのであつて、殊に社会的集団現象の生起の基盤である人口及びその社会的、産業的、地域的構成に就て悉皆的な統計的智識を必要とし、従つてそれを得るためには悉皆調査がなされねばならないのである。故に今日の国勢調査の如き悉皆調査は、只にそれ自体の目的（人口の自然的、社会的状態の静態調査）のみならず、又任意標本調査の実行に必要な基礎資料の獲得をも目的として調査の設計並に集計、整理がなされているのである。しかしこれ等の条件が整いさえすれば任意標本調査は悉皆調査に劣らず、時には悉皆調査以上に正確な調査方法であるため、社会統計調査に於ては悉皆調査は専ら基本的事

項のみに止めて複雑な事項は任意標本調査に委ね、或は又定期的になされる悉皆調査の中間的、補完的調査に任意標本調査を利用するという方向に發展しつつあるのである。

かくて社会的集団現象の正確な統計的認識の可能性は、悉皆調査と任意標本調査との併用により若しくその範囲を拡大し得たのであるが、尙未だ任意標本調査の条件が整わず又本質的に悉皆調査は勿論標本調査も不可能な現象が存在するため、旧来の標本調査法並に推算は社会統計調査に於ては不可欠の方法である。

以上により社会統計調査に於てとられる諸種の方法特に悉皆調査と標本調査との関聯を、發展的に概観する事によつてそこに於ける各々の地位を明らかならしめたのであるが、次は各方法に就てその方法的特質を考察して行こう。

## 二、統計調査法（統計的方法）

統計的方法による社会的集団現象の調査、研究は、

社会的集団現象を構成する個々の要素の個別観察、分類、集計という手續過程を経て行われるのであつて、このために社会的集団現象の数量的な精密な認識が可能となるのであり、そして統計的把握が数量的に行われるために、数学的思考方法により数理的解析手續を加えることによつて、社会的集団現象の性質を研究することが出来るのである。以下統計的方法による社会的集団現象の認識はどのようにして行われるのであるかを明らかならしめよう。

社会的集団現象を統計的に把握するためには、先づ「統計集団 *statistische Masse*」を構成しなければならぬ。即ち統計的方法は社会現象を一個別現象として、思惟により定性的に把握するのでなく、それを集団現象として且集団構成要素の個別観察に基いて、数量的に把握する方法であるため、先づ社会的集団現象の事物的性質を集団表章とし、それに存在の本質的条件たる時間的、場所的規定を附加することによつて、集団構成要素（統計単位 *statistische Einheit*）の数を限定し

て一定の時所に於て存在する集団たらしめ、次に社会的集団現象に就て明らかにせんとする諸性質を此の集団の性質として規定しなければならぬのである。かくて始めて社会的集団現象は統計的方法の対象となり得るのであつて、故にこれを統計集団という。

しかしながら理論的正確性に於て構成せられた統計集団は、その構成要素の個別観察を可能ならしめねばならないために、統計単位の実際観察が可能であり且その存在を明確に限定し得るが如き時間的、場所的規定に変更し、又統計集団の性質も個々の統計單位に就てその所有の状態（質的性質に就ては所有の有無、量的性質に就ては発現の程度）を、正確に観察し得るが如き具体的規定に改められねばならないのである。

統計集団の性質のうち集団の性質の種類を表す如き概念を「集団性」、その種類を集団性の「方向」といい、集団性を上位概念としそれによつて総括されるより具体的な諸性質を「標識 (Markale)」と云う。例えば人口なる社会的集団現象の有する一性質である「体性」は集団性であり、

## 統計調査法

「男」及び「女」はそれに従属する標識である。

この場合根本的には統計集団の理論的構成、即ち客観的存在たる社会的集団現象の理論的把握に於ける調査者の観点の妥的性が問題であるが、尙更に方法的にはその具体的規定への改革は、多くの場合理論的統計集団と具体的統計集団との間に乖離を生ぜしめる事となり、かくて統計的方法は出発点に於て事実を歪曲したものを対象として持つ事となり易いのである。例えば工業平均賃銀を求めるときは合理的統計集団としてはすべての工場の労働者を以て構成されるべきであるが、統計的観察の可能性の点より従業員三十人以上の比較的大きな工場の労働者のみを以て具体的統計集団を構成するが如きこれである。この時得られた平均賃銀は労働条件や給与の劣悪な小工場の実状を反映していないため、本当の工業平均賃銀よりも若干高い値であると考えられるのである。茲に統計の信頼性（階級性、欺瞞性）の問題が発生する根拠が存在するのであつて、統計の利用に際して最も注意すべき点である。<sup>1)</sup>

かくて統計的方法の対象が決定すると、次に統計集團の有する諸性質（各方向の集團性）を個々の統計單位に就て観察するのである。此の個別観察に於て統計單位は、先づ一方方向の集團性に就てその有する具体的性質を観察され、その性質に應じて標識が与えられるのであり、このようにして一個の統計單位は順次集團性の方向丈その有する性質を観察され、それぞれ適当な標識を与えられるのである。

社会的集團現象の統計的觀察に於ては、統計單位の個別觀察は当該現象に關聯を有する主体（個人、団体）を通じてのみなされ得るのであるから、統計集團の性質、其他統計單位に就て觀察せんとする事項を主要内容とする調査票を設定し、それを個々の被調査者に配付して調査事項に適宜記入の上報告を得る事によつて統計單位の觀察が行われるのである。しかしながら被調査者と調査者とは一定の社会的、政治的、經濟的關係に於て存在するものなるため、調査事項が經濟的、道德的事柄の如き被調査者の社会生活に重大な影響を

及ぼすものなる場合は、その報告が正しいか否かが問題となるのであり、茲に統計の正確性の吟味の最大の根拠が存するのである。

このような個別觀察をすべての統計單位に就て行くと、次に標識毎にそれを有する統計單位を分類しその数を合計する事によつて、各標識を同じ性質の單位の数によつて具体的に規定するのである（かくして得られた標識の數量的規定を「統計値」といい、個々の統計値を総稱して「統計」という。）

この統計的規定を受けた諸標識を集團性によつて統括する時統計集團の大小が明らかになり、且それは標識を同じくする統計單位の小集團即ち標識を集團表章とする部分集團に分割されてその構造を明らかにし、それ等の部分集團相互の乃至は部分集團の統計集團全体に對する數量的關係によつて、一方方向に於ける統計集團の性質が數量的に規定されるのである。このような統計集團の構造分析によるその性質の規定は集團性の方向毎に行われ、更に諸方向の集團性を組合せる事

によつて、一方向の集団性に就て明らかにされた統計集団の構造は更に他の方向の集団性に關して構造的分析せられて、その性質の詳細な數量的規定がなされるのである。最後にすべての方向の集団性が集団表章によつて綜合される事によつて、統計集団の諸種の性質が全体的關係に於て數量的に明らかにされるのである。以上のような統計集団の分析は概念による分析ともいふべく、現象の性質相互の論理的、事實的關聯を基礎にして集団の數量的な分析が行われるのである。

かくて統計集団の性質は部分集団相互の又は部分集団と統計集団全体との數量的關係として表されているのであるが、その相互關係の概観は容易でないため統計集団の性質の簡約な表章が必要となるのである。そしてそれは集団性が質的な性質の場合には、各標識を有する統計單位の存在の割合即ち比例數を求める事になり、若しそれが量的な性質の場合には、すべての統計單位の有する標識即ち數値の平均値、及び平均値より各標識(數量値)の偏倚する程度(分散)又は標

準偏差)等によつて測定する)、或は又各數量値を有する單位の數の變化の状態即ち度數分布の形態を數學式(分布函數)によつて表現し、更にその特徴を示す値——歪度、尖峰度等を求める事によつて可能となるのである。

尙統計の數量的性格は更に進んで複雑な數理的解析手續を施す事によつて統計集団の性質を解明する事を可能ならしめるのである。これ等の統計集団の分析は數學的論理に導かれて行われるのであるため、數理による分析という事が出来る。

以上のようにして社会的集団現象の大いさ並に性質を數量的に精密に調査、分析する統計的方法は、その方法的手統過程の特質により統計集団の觀察、記述の過程と、その記述結果に數理的な解析手續を加える過程とに分ける事が出来る。前者は大量觀察の過程、後者は統計解析の過程といわれ、その方法的規定を夫々「大量觀察法 Massenbeobachtungsmethode」「統計解析法 method of statistical analysis」といふ。そして

大量観察の成果たる統計を「基礎統計数」といふ、統計解析の結果たる統計を「誘導統計数」といふ。

### 間接大量観察法

尙大量観察の技術的手統過程の基礎である集団構成要素の個別観察が、統計調査夫自体のためになされるのではなく、他の目的から調査、記録せる結果、例えば行政事務執行に伴つて蒐集した事実の記録や社会等の民間各種事業の記録を以て代用し、それより統計を作成し得る場合がある。例えば出生、死亡、婚姻等の人口動態統計は何れも行政官庁への届出の結果によるものであり、尙此の外に貿易統計、金融統計、所得統計等多くの統計がこれによつて作成されているのである。

この方法は統計資料の蒐集が間接的になされるのであるため「間接大量観察法」といふ、これによる調査を「間接調査」又は「第二義統計調査」、これによつて得られた統計を「第二義統計 sekundäre Statistik」

という。（従つて集団構成要素の個別観察の手統を経て行ふ調査を「直接調査」、それより得られる統計を「第一義統計 Primäre Statistik」といふ。）

間接大量観察法による時は利用する記録が原調査によつて絶対的な限定を受けているのであつて、必ずしも統計集団の規定と一致するものではなく、又完全に集団構成要素が観察されているかは疑問であるため、これによつて得られた第二義統計の価値は自から第一義統計よりも劣るのである。しかしながら此の方法では多穴の労力と費用を必要とする大量観察の個別観察、分類、集計の技術的過程が著しく省略される利点があり、原調査に於てその結果の統計作成への利用を考慮して調査、記録を行ふ事によつて正確な第二義統計の獲得が可能であるため、社会的集団現象の調査に於て極めて重要な役割を果たしているのである。

①② 統計の信頼性、正確性に就ては左の諸書に詳しく論ぜられてゐる。

嵯川虎三「統計学概論」第二章第二節大量観察法 「統計  
利用に於ける基本問題」第二章大量観察に於ける理論と技  
術

杉栄「理論統計学研究」第四章及び第五章参照

### 三、悉皆統計調査の正確性

統計的方法による社会的集団現象の認識は統計集団  
を媒介としてのみ行われ、茲に統計の信頼性の問題が  
発生する根拠があるのであつた。そして統計集団の観  
察は集団構成要素の個別観察、分類、集計なる技術的  
手続を経て行はれるのであるが、この時構成要素のす  
べてを洩れなく観察する事によつて始めて、統計集団  
の大きさ並に構造の正確な認識が可能となるのであつ  
て、若し一部の要素を観察対象より逸するならばそれ  
の正確な認識は得られないであらう。従つて社会的集  
団現象の事実としての大きさ並に構造の絶対数による  
記述即ちその統計調査が目的である場合には、集団  
構成要素の悉くを数え上げるのでなければ誤差のない  
正確な統計は得られないのである。かくて社会的集

団現象の調査方法に於ては、悉皆観察 *erschöpfende Beobachtung* が原則となるのである。

従来ドイツ社会統計学派に於ては大量観察法、従つてそ  
れを包摂する統計的方法は悉皆観察をその本質的条件とし、  
悉皆性を缺くものはその代用法——大量観察代用法又は統  
計的方法の代用法——とされて来たのであるが、それは社  
会的集団現象の統計調査に於ては悉皆観察によつて始めて  
統計の正確性が保証されるからであつた。しかしながらそ  
の事より方法としての大量観察法が悉皆性を本質とすべき  
であるとは考えられないために、我々は大量観察法はその  
語義通り前節に述べた如き方法手続による集団 *Masse* の  
観察方法であつて、悉皆観察たると一部観察たるを問は  
ないものと解さんとするのである。

茲に「社会的集団現象の調査方法に於ては」とは、  
若し社会的集団現象の一般的性質、規律性の研究が目  
的である場合には、必ずしも統計集団構成要素の悉皆  
観察が統計の正確性を保証する所以ではないからであ  
る。蓋し、社会的集団現象の一般的性質、規律性を研  
究せんとする場合、統計的方法の対象たる統計集団は

その社会的集団現象の特定の時所に於ける状態を表すに過ぎないため、假令此の統計集団を悉皆観察するも矢張りその社会的集団現象の一部分の観察に過ぎず、故にこの現象の一般性、規律性の認識の正確性を絶対に保証するものではないのである。大體統計的方法による社会的集団現象の一般性、規律性の認識の可能性の根拠は大数法則の作用によるのであつて、大数法則が作用するのに十分な大多數の現象を集団的に観察すればよいのである。そしてこのように多数の現象の集団的觀察の可能な極限が統計集団の全部觀察であるのであつて、従つてこの場合には統計集団の悉皆觀察は必ずしも必要ではないのである。

そして社会的集団現象の一般性、規律性の認識は平均値、比例数等の誘導統計値によつてのみ可能であつて、その絶対的な大いさを表す基礎統計数によつては不可能な場合が多い。それは同じ社会的集団現象と雖も時所を異にするに従つてその大いさは必ずしも同一ではないからである。このように社会的集団現象を絶

對値に於て問題にするのではないならば、悉皆觀察の結果に略等しい結果を一部分のみの觀察によつても得る事が出来るのであつて、この事は数学上の大数法則によつて数理的に証明せられるのである。故に特定の時所に於ける社会的集団現象の絶対的な大いさ並に構造の把握が目的である場合には、一部觀察によつては誤差のない値は決して得られないのであつて従つて悉皆觀察は本質的な条件となるのであるが、社会的集団現象の事實的認識が平均値、比例数等によつて行われる限りに於ては、統計調査は部分觀察の方法による事が出来るのであつて、この事は後節標本調査法に於て考察されるであらう。

そして社会的集団現象の統計調査の實際に於ては、悉皆觀察は大量觀察の技術的過程に於ける誤差の發生を激化し、従つて必しも正確な結果が得られるとは限らないのであり、そして悉皆觀察の実行は多大の費用、労力及び日時を必要とするため複雑な調査は困難であり、これ等の犠牲との關係に於て（悉皆）統計調査の

相對的に不可能な現象が多く存在するのである。しかしそのような現象に就ての統計的認識の必要性は大であるため、假令理論的には誤差が存在するにしても統計調査の容易な非悉皆的方法による調査が行われるのである。この場合は注意すべきは假令誤差のある統計と雖も、その統計の利用目的に鑑みてその程度の誤差は許容し得る限りに於ては十分正確であると云えるのであつて、統計の正確性は誤差の有無のみによるものではないという事であるが、それはやがて明らかにされるであらう。

悉皆調査は原理的には誤差を伴わない正確な統計を獲得せしめるのであるが、実際上は必ずしもそうではないのである。しからばそれは何によるのであるか、その根源を正確に認識する事は、悉皆調査の誤差を減少せしめてより正確な統計を作成するために必要な事柄である。故に次にそれを考察しよう。

大体統計の誤差は、社会的集團現象の統計集團への規定に際しての兩者の乖離に於て統計の信頼性として

問題になり、次に統計集團構成要素の個別觀察、分類、集計の技術的手続の實行に伴つて發生する誤謬に就て統計の正確性として問題にされるのである。このように大量觀察法はそれ自体のうちに統計の誤差を生ぜしめる根拠を有するのであるが、それが悉皆觀察の場合には如何に現れるかを考察すればよいのである。

社会的集團現象は広大な地域に亘つて非常に多くの要素より構成されて存在し、その統計的觀察は当該現象に關聯を有する主体を通じて調査票によつてなされるのであるため、多くの調査員の協働とそれを効果的ならしめ調査票の円滑な運用を可能ならしめる統計調査機關の組織の整備、發達を必要とするのである。従つて社会的集團現象の悉皆統計調査の實行は多数の人員と多額の費用を必要とし、故に構成要素の個別觀察に多くの手数を要するようた複雑な事項の調査は、

假令理論的には悉皆觀察が可能であるにしても、それに要する犠牲のために實際上は到底不可能である。そして大量觀察法の直接の対象である統計集團は、社会

的集団現象を統計的観察即ち集団構成要素の個別観察を可能ならしめるように規定したものであるが、此の個別的観察の可能性は個々の統計調査に於ては只単に技術的な可能性のみならず、更に実際の個別観察手続の遂行に要する労力、費用との關聯に於て問題とされねばならないのである。従つて統計集団は社会的集団現象の調査の目的に鑑みて、その社会に於ける個別観察技術の發達の程度を基礎にして、与えられた労力と費用とにより悉皆観察の可能なように構成されるのであつて、かくて悉皆統計調査に用い得る労力、費用が十分でないならば社会的集団現象と統計集団との乖離が著しくなり、統計の信頼性を害う可能性が増大するのである。

そして統計集団の實際的観察の過程即ち大量観察の技術的手続過程は、多くの調査員の協働の下に、調査票による調査主体と一定の社会的、政治的、経済的關係に於てある被調査者よりの報告としてなされるのであるから、その事より種々の誤差を生ぜしめるのであ

る。即ち前節に於て述べたように調査事項によつては報告内容の真偽が問題とされねばならない事は勿論であるが、このような意識的な誤謬の外に尙調査票への記入に際しての無意識的な誤記、脱漏の發生が予想されるのであり、又問題によつては調査員の側よりする被調査者への影響が誤れる報告を生ぜしめる恐れが多分に存在するのである。これ等の観察の誤謬特に最後の誤りの如きは、調査票の表現形式を工夫したり調査員を訓練する事によつて或る程度回避する事が可能であるが、悉皆観察の場合のように被調査者が広範囲に極めて多数存在し、従つて多くの調査員を動員しなければならぬ時には、調査員の質も不揃になりその教育も十分に行われず、又調査員の被調査者の報告に対する注意、指導も行届かないためにこれ等の観察の誤謬は極めて大となるのである。そして又個別観察結果の分類、集計に於ても、如何に正確にそれを行わんとしても老大な資料の分類、計算には誤謬が当然に發生するであらう。

かくて大量観察の技術的過程に於ては多くの誤差が發生する可能性が存在するのであるが、それ等のうち正確ならんとしてつても無意識的に發生する誤謬の大部分（例えば誤記、計算間違等の如きもの）、は観察対象の数が多くなるに従つて誤差法則に基いて相殺されると考えられるため、悉皆観察の結果に於ては略々消失すると看做す事が出来るが、中には不正申告の如き意識的誤謬と同様に観察結果の綜合に於て、誤差として一定の bias 偏りを生ぜしめるものがあるのであつて（例えば調査員が被調査者に与える影響や次に述べる申告洩れ等）それ等による誤差は観察対象の数の増大と共に益々激しくなり悉皆調査に於て極限に達するのである。

そして観察さるべき対象は非常に多く存在するのであるから完全に凡てを握える事は決して容易な事柄ではなく、被調査の脱落、統計單位の観察洩れは当然に予想すべきであり、尙更に被調査者の都合により調査票の報告が得られず、或は又著しく遅延したため集計

に間に合わない事は統計調査の實際に於ては必ず生ずるものであつて、この無回答者の追索は極めて困難な事柄である。そしてこれ等の観察の欠除は観察対象の数が増大するに従つてより大となるため、悉皆調査と雖も完全に悉皆性が確保される事は殆どないのであつて、この点よりしても悉皆観察の結果は誤差を有するのである。

そして個別観察の結果の分類、集計手続の実行には多くの人手と相当の日時を必要とするのであつて、統計調査が悉皆観察による時はその老大な観察結果を整理して統計を獲得する迄には極めて長い期間を経なければならぬのであり、統計機械の發達はそれを著しく短縮せしめたとはいへ尙此の欠点を除去する迄には至つていないのである。従つて社会的集團現象が變化の激しい現象である場合には、悉皆統計調査が十分の勞力、費用を以てし又如何に大量観察の技術的過程が正確に行われたとしても、統計が得られた時には既にその値は單なる過去の事実を物語るものとしての価値

しかなく、現在の正確な状態を知る手掛りとはなり得ない事が多いのである。このように悉皆調査は時の遅れによる統計の誤差を大ならしめるために、實際生活の指針を得るための統計の利用の上からは必ずしも正確な方法ではないのである。

以上により明らか如く大量観察の技術的手続過程に於ては必然的に誤差が発生するのであるが、これ等の誤差の大きさに就てはそれを測る客観的な尺度は存在しないのである。熟練せる統計家によれば或る程度それを推定する事は可能であるが、その根拠を正確に示す事は不可能であり主観的な評価に負うところが極めて大である。

そしてそれ等の誤差は理論的には誤差を有しない筈の悉皆観察によつて却つて増大せしめられるのであつて、むしろ理論的には誤差を伴う非悉皆的な一部観察の方法による時は減少せしめる事が出来ると考えられるのである。このように悉皆調査によつても必ずしも誤差のない値が得られるものではないとすると、利用

目的よりして許容し得る範囲の誤差の統計であるならばそれは正確であるといえるであろう。従つて統計の正確性は只単に得られた数字の誤差の絶対的な大きさのみならず、その利用価値とそれの獲得に要する犠牲とを併せ考慮して、これ等の綜合に於て考えられねばならない問題である事が判るのである。

かくて悉皆統計調査は実際上は必ずしも社会的集団現象の正確な統計を作成するものではないのであつて、その程度を測る手段のない誤差を含む値が得られるに過ぎないのである。しかしながら以上の大量観察に於て発生する誤差は、調査事項が非常に簡單、明白な事柄であり且多大の労力、費用を用い得る時は極度に減少せしめる事が可能であるため、悉皆統計調査は社会的集団現象の大きさ並に構造に就ての絶対的に正確な値を与えるであろう。今日センサス (Census) として悉皆調査が行われているのは主としてこのような場合である。

① センサスは元來人口動態の直接的悉皆調査即ち國勢調査

を意味するのであるが、今日では人口静態の外に特定特点に於ける諸種の社会現象の悉皆調査をもセンサスというようになった(例えば農業センサス、工業センサス、商業センサス等)。しかし唯単にセンサスといえば同勢調査をいうのである。

#### 四、其他の統計調査法

我々の認識の目的が社会的集團現象の事実としての大いさ並に構造の把握にある時は最も正確な方法は悉皆調査であるが、これは多くの労力、費用並に日時を必要とするため実行が困難であり、且実際上は必ずしも正確な結果が得られるとは限らないのであつた。しかし社会的、経済的生活の實踐に於ける社会的集團現象の実情の把握の重要性は、絶対的に正確な値ではなく多少の誤差を有するとも、容易に迅速且低廉に又信頼性の大なる統計の獲得を必要とするのである。此のためにとられる方法には「推算 Schätzung, estimation」と「標本調査 typische Beobachtung (一部調査 Teilerhebung)」がある。<sup>註</sup>

#### 統計調査法

推算は既知の統計より數理的、計算的に新しい統計を作成する方法であつて、集團構成要素の個別觀察の方法手続を含まないために統計解析法的方法といふ事が出来、標本調査(一部調査)は社会的集團現象の一部分を大量觀察手続によつて集團的に觀察し、統計を獲得する方法であるために多分に大量觀察法的方法であつて、故に両者はその方法的性質を異にするのである。従つて推算は(悉皆觀察のみならず仮令一部觀察であつても)大量觀察法によつて把握する事の不可能な、又は困難な社会的集團現象に就て極めて有用な統計的認識の方法であり、標本調査(一部調査)は大量觀察法によつて把握し得るが、悉皆觀察によるため相對的に不可能乃至は困難な場合に重要な統計調査の方法である。

<sup>註</sup> アンケート Enquete (訪問調査)

尙從來はこれ等の外にアンケートが挙げられるのが普通である。しかしながらアンケートは社会的集團現象の數量的な把握を目的とするのではないのであつて、社会的集團現象の性質を明らかにするためにそれと特別の利害關係を

有する者乃至はその方面の専門的智識を有する者に、一定の事項に就てその知るところを問合せる方法である。アンケートによつては社会的集團現象又はその部分集團現象の大小を示す統計を得る事が出来ないため、これは大量観察法に於て又統計的方法の一種と看做するは出来ないものであつて、その結果は社会的集團現象の理論的把握のための一資料として有用であるため、大量観察の補完的な方法と見るべきである。

## (イ) 推 算

推算は或る社会的集團現象の大きさ又はその部分集團現象の大きさを、既知のその集團現象を表す統計、又は当該集團現象と密接な関係を有する他の社会的集團現象を表す統計を利用して、未知の集團と既知の集團との間の一定の具体的な関係を辿つて計算的に求める方法である。このような推算の可能性は、社会的集團現象は多くの他の現象と相互に関係を有し影響し合ひ乍ら存在し生成、発展しつつあるものなるため、直接社会的集團現象を観察しなくとも、それと密接な関

聯を有する諸要因の大きさ及びその変動状態を正確に知るならば、双方の事実的關係を辿る事によつてその社会的集團現象の大きさを推定する事は可能であろうし、又社会的集團現象夫自体の發展は連続的に行われる場合が多いために、その大きさの既知の動きより未知の大きさを推定する事が出来るであらう事によるのである。

推算の推理方式即ち計算方式は、統計的に既知の社会的集團と今知らんとする夫との間の実際的な關係によつて規定されねばならないのであるが、それは当該現象に就ての科学的智識による兩集團現象間の理論的關係に基いて決定されるか、又は過去の統計的關係を示す値や方程式によつて与えられるのである。例えば一定時点に於ける靜態的な人口現象の悉皆大量観察である國勢調査は、どの國に於ても五年乃至十年毎に行われるため、その中間年度の人口は推算によつて求められねばならないのである(推計人口)。此の時人口理論に基いて、實際の國勢調査人口に年間其後の出

生、死亡の差及び移出入の差を加減して期末人口を推算するという推理方式は前者の例であり、又は数理的に両調査年度の調査人口よりその間の年平均増加率を（幾何平均的に）求め、その増加率を用いて中間年度の人口を推算する方式は後者の例である<sup>1)</sup>。

かくして決定される両集団現象間の関係は統計集団間の関係であらねばならない事は、その推理方式より決る計算方式には統計集団を表す統計値が当嵌められるのである事よりして明らかであろう。

推算は両集団間の理論的乃至は統計的關係によつて決定された一定の計算方式に、既得の統計値を代入して計算を行う事によつて統計を求める方法であるため、その結果の正確性は、先ず算定の基礎となる統計値が正確なるものである事と、次に未知の集団現象と統計的に既知の統計集団との間の関係の規定、即ち計算方式が正しい事に依存するのである。

このうち後者の条件は極めて重要であつて、推算が往々にして「数字の独り歩き」に陥り「数字の遊戯」

となる危険性の根拠は茲に存するのである。推理方式が密接な關聯を有する現象のすべてを統計集団間の關係として、且實際によく妥当せる關係に於て考慮に取入れたものであるならば、それによる計算方式は正確なものとなるであろう。しかしながら多くの場合はその様な必要な關聯のすべてを考慮する事は不可能であり、又その計算に必要な統計で利用し得ないものが多いあるため、却つて複雑な計算方式程その結果は危険であるのが実情である。又推理方式が過去の統計的な規律的關係に基く時は、その關係は單に過去の事實に於て一般的に見られた結果であるに過ぎないのであつて、それ以外の場合に於ても絶対的に妥当するその現象の普遍的な數量的關係と看做し得る實質的な基礎が薄弱であるため、この推理方式の価値は理論的推算よりもむしろ低いといわねばならぬ。故に推算の結果には嚴密な正確性は決して望めないであつて、假令正確な値といつてもそれは概數としてのみい得るに過ぎないのであり、又その誤差の測定は到底不可能であ

る。

又推算の基礎になる統計値が正確でなければならぬ事はいう迄もないが、若しそれに直接調査によつて得られた値でなく推算によつて得られたものを利用する場合、特に同じ計算方式によつて推算を繰返して行く時は、一定の傾向の組織的な推算の誤差が累積されて行くために、適当な時期に直接調査を行つて基礎資料を正確ならしめねばならないのである。

このように推算は正確な値を得るものではないが、大量観察法によらずに社会的集団現象の統計的推定を可能ならしめる方法であるため、社会的集団現象の将来の状態の予想（又は過去の状態の推定）にも応用する事が出来るのである。尤もこのような将来の予想は、多くの場合現在の条件が将来に於ても不変である事を前提として始めて可能となるのであるが、一般にそのような保障は得られないためその結果の正確性に就ては極めて薄弱である。

社会現象の統計的研究は最初推算から始つたのであ

つて、その後統計調査機関並に制度の整備充実と共に大量観察法による正確な統計が作成されるようになってのである。しかしながら今尙推算の重要性は決して減少してないのであつて、人口の推計、国民所得の算定を始めとして其他経済方面に於て広汎に用いられているのである。

- ① 推算の諸方式に就ては、森田優三「統計学汎論」二五六—二八〇頁に詳細に論ぜられている。

### (ロ) 標本調査法（一部調査法）

一部調査法は統計集団構成要素の凡てを観察する事なく、それより適宜選択した一部分の構成要素のみを集団的に観察して、悉皆観察の結果得られるであろう統計集団の構造を推定せんとする方法である。

この時直接統計的観察の施される統計単位の集団を「標本、試料 sample」とし、従つて一部調査法は又標本調査法ともいふのである。そして標本を抽出する母体たる集団を「母集団」といふ。しかし標本と母集団とが明瞭に區別されるようになったのは、任意抽出法による標本調査法が

發達するようになってからであつて、且その場合には特別の意義が与えられているのである。(任意抽出法の項参照)

しからば標本調査はどのようにして統計集団の構造を推定する事が出来るのであろうか。第三節に於て述べたように事実としての社会的集団現象の把握に於ても、その絶対的な大いさは必ずしも問題でなく平均値、比例数等の統計集団の構造の簡約表章手段による表示が要求される場合には、必ずしも統計集団の悉皆観察は必要ではないのであつて、そのうちの一部分のみの集団的観察によつても十分正確な値が得られるのである。それは平均値、比例数等は統計集団の数量的構造、即ち統計集団全体の大いさに対する各部分集団の大いさの数量的相互關係を表す値であるため、統計集団より一部分の構成要素を抽出し構成した集団即ち標本が、統計集団の悉皆観察によつて明らかにされるであろう数量的な構造と同様の数量的構造を有するよ

得られる平均値、比例数等は略々悉皆観察の結果に等しくなるであらうからである。

かくて標本構成要素の選択さえ適當になされるならば、標本調査によつてもそれより得られた平均値、比例数等によつて悉皆調査の結果を推定する事が出来るのである。しかしながら此の場合実際には悉皆調査の結果は判つていないのであるから——そうであるからこそ標本調査によつてそれを推定せんとするのである——統計集団の縮図となり得るような標本を作る事は容易でなく、従つて、標本調査の結果が悉皆調査の結果に完全に一致する事は到底期待し得ないのである。故に標本調査による時は当然誤差を伴うのであつてそれ文正確性を欠き、若し誤差のない値が必要であるならば悉皆調査によらねばならないのである。

うちに、即ち統計集団の縮図となるような標本が構成されるならば、標本の集団的観察による構造分析の結果

しかしながら第三節に於て明らかならしめた悉皆調査の欠点——大量観察の理論的過程に於ける信頼性の毀損、技術的過程に於ける観察誤差の増大、多大の費用と労力及び時の遅れによる誤差の激化等——は標本

調査による時は著しく緩和されるのである。蓋し標本調査法はその方法的性質に於て悉皆調査と同様大量観察法によるものであるため、大量観察に於ける誤差を免れる事は出来ないであらう。しかしながら標本調査は観察すべき統計単位の数を大幅に減少せしめる事が出来るために、社会的集団現象と統計集団との乖離を少くして統計の信頼性を高め、又単位の個別観察も十分な注意を以て行われ無回答者の追求も徹底的になし得るため観察誤差は減少せしめる事が出来る。その上統計調査に要する費用、労力は少くして済み調査が簡便に行えるために社会的集団現象の統計調査の範囲は拡大し、且観察結果を統計に作り上げる事も短時間で出来るため統計の利用価値を著しく増大せしめるのである。かくて標本調査は方法夫自体に於て誤差を伴うのであるが、統計調査の實際に於ては標本の構成が適切になされるならば必ずしも不正確な方法とはいえないのであつて、むしろ悉皆調査よりも正確な統計が獲得される場合が多くあるのである。

標本調査法は標本構成要素の選択方法の相異により「有意選択法 method of purposive selection」と「任意抽出法 random sampling method」とに分ける事が出来るのであつて、前者は観察対象を何等かの基準に照らして意識的に選択する方法であり、後者はそのような主観的意図を加えずに全く偶然的に観察対象を抽出する方法である。次に先ず有意選択法による標本調査から考察して行こう。

### (1) 有意選択法

有意選択法は統計集団の構成要素のうちより、統計集団の調査せんとする性質と極めて密接な關聯を有する諸性質（これを「対照 controls」という）に就て、すべての構成要素の典型的乃至は平均的な要素と看做し得るものを選んで標本を構成する方法である。

統計集団の一部分のみの観察によつて統計集団全体の性質を知らんとする場合、統計集団を構成する要素のうち調査目的たる性質（これを「目印 characters」

と呼ぼう)に就て典型的乃至は平均的な要素を選出して標本を構成する事が出来るならば、その標本は統計集団を代表しそれより得られた値は統計集団の性質を表すものと看做す事が出来るであらう。蓋し、標本の構成要素は統計集団の構成要素のすべての典型乃至は平均であるからである。しかしながらそれでは如何なる集団構成要素が典型的、平均的であるかは、統計集団の悉皆観察の結果を俟つて始めていい得るところであつて、その前には決して正確に規定する事は出来ない問題であり、かくて標本調査はその結果を前提に必要とするという矛盾に陥るのである。しかし今調査せんとする性質即ち目印と密接な関聯を有し、相互に影響し合うが如き諸種の性質即ち対照に就て、すべての統計単位の典型乃至は平均であると看做し得る要素を選挙するならば、それは目印に就ても又すべての要素の典型乃至は平均と考える事が出来るであらう。従つて対照に關して典型的、平均的な要素を以て構成せられた標本は、同時に目印に就ても又典型的、平均的な

要素によつて構成せられてゐると看做す事が出来るであらう。若し統計集団が極めて異質の要素より成つてゐる時は、それ等全体を通じての典型的、平均的要素を求めざる事は不可能に近いであらう。従つて此の場合には対照によつて統計集団を同質的な要素より成る部分集団に分割し、各部分集団に於て典型的、平均的要素を選挙して標本を構成しなければならぬ。例えば一村の米の収穫高を坪刈調査によつて推定せんとする場合の全耕地を水系によつて類別し、各水系毎に作柄の中庸と思われる田を一枚宛坪刈して水系毎の平均反収を決定し、それを加重平均して村全体の平均反当り収穫量を決定するのである。

標本を構成する場合統計集団より選出せられる要素は、統計集団が小さい時は個々の統計単位に就て代表的なものを選ぶ事が可能であるが、統計集団が極めて大きい時はそれは困難であるため、便宜上統計単位の集団(部分集団)特に地域的集団(市、郡、町、村等)に就て典型的、平均的なものを選定する方法がとられるのである。

以上により明らかかなように有意選挙法は対照によつ

て間接的に代表的な標本を構成せんとする方法であるため、これによつて標本調査を行うためには対照たり得る如き統素集団の性質に就ての智識の存在を必要とするのである。しかし調査対象たる社会的集団現象に就て何等の智識も持つて居らず、これから行う統計調査によつて始めてそれを獲得するという場合は殆どないのであつて、大抵はその現象に就て科学的な又は常識的な智識を有するのであり、時には統計的な智識——その現象に就ての過去の統計調査又は類似の調査によつて得られた——を持つている事もあるのである。故にこれ等の智識を綜合して調査目的たる性質と特に密接に關聯し相互に影響し合うものを対照とし、それに就て典型的、平均的要素を決定するのである。そしてこの場合対照に就て統計的な智識が得られているならば、構成要素の選択は只単に概念的、典型的、平均的と看做されるものが選ばれる文ではなく、統計的な値に於て平均的なものが、且標本に於ける対照の平均が統計集団に於けるそれと一致するように選出され、

対照に就ての平均性が統計的に明らかとなるのである。例えば農業経営に就て標本調査を行わんとする場合、全国の村を地理的特徴、耕作物内容、他産業との關聯等によつて類別し、その各々より一定数宛抽出して標本を構成する時、標本に於ける対照に就て求めた平均値、比例数等が全国に於けるそれに一致するように、標本構成要素たる村を選択するのである。

この場合標本構成要素の数即ち標本の大小さは如何に決定さるべきであらうか。若し統計集団の構成要素のすべてに対して厳密に典型的、平均的な要素が存在しそれが選定し得るならば、一個の要素の觀察によつても統計集団の性質を知る事が出来るであらう。しかしながら社会的集団現象は自然現象と異り、それを構成する個々の現象は典型的現象ではなく極めて非典型的な現象であるため、完全に典型的な平均的な要素が存在し選出し得る事は殆どなく、その典型性、平均性は大体に於てのものであるに過ぎず、又それも一二の対照に就てならば兎に角もすべての対照に就て典型的、

平均的な要素を見出す事は容易でなく、更に仮令対照に就て同質的な要素が選出されたとしても、それ等は目印に就ては非典型的であると考えられるから、比較的多数の要素を以て標本を構成して、各標本構成要素のこれ等の非典型性、非平均性を相殺、消去しなければならぬのである。従つて標本はこの相殺を可能にして（即ち大数法則が作用して）標本の代表性を高め得るような大いさであらねばならないのであるが、その数の如何を具体的に決定する客観的な基準は容易に得られないのである。従つて実際上は標本調査の費用の制約の下に、略々これ位とれば非典型性を消去し得るであらうと考えられる程度の大いさに決定されているのである。

以上により明らかな如く有意選択法による標本調査の正確性の基礎は、対照に就て典型的、平均的なるため目印に就ても又典型的、平均的であらうという仮定にあるのであつて、従つて対照の数が多く且それと目印との關聯が密接である程標本調査の結果は正確であ

るといえるのであらう。しかしながら対照に就ては典型的、平均的な要素（従つて標本）と雖も、その事より直ちに目印に就ても然りであるという事は必ず主張し得ないのである。成程概念的には、対照と目印との間に密接な關聯が存在する限り、前者に就いて典型的であるならば後者に於ても又典型的であるといえるであらう。しかしながらその時その標本より得られた目印に就ての値（平均値、比例数）が統計集団のそれに等しいという事は、社会的集團現象を構成する要素的現象（統計單位）の非典型性の故に一般的には主張し得ないのである。若し対照に就て正確な統計的知識を持つて居り、統計的に平均的な標本が構成され且対照と目印との間に極めて密接な相關々係がある事が判つてゐる時は、その標本は目印に就ても平均的となるであらう。しかし實際は対照に就てそのような詳細な知識が得られてゐる場合は殆どなく、仮令あつても時の経過による誤差の存在が予想される事が多いために、目印に就て統計的に平均的な標本を構成する事は極めて

困難である。かくて有意選択法によつては概念的典型性は確保する事が出来るが、統計的（数量的）平均性は必ずしも得られないのである。しかしながら標本調査法の正確性は標本の観察によつて得られた値が、統計集団の悉皆観察によつて得られるであろう値に一致するや否やによつて決定されるのであり、従つて標本の代表性も此の点より評価されねばならないとするならば、しからば有意選択法は一般に代表的な標本の構成を保証し得るものではないといえるのである。

その上既得の智識を利用して典型的、平均的な要素の如何を、多くの統計集団構成要素中より選択し得るように具体的、一義的に規定する事は、社会現象の非類型性の故に極めて困難な事柄であつて調査計画者の創意に負うところが穴であり、故に決定された典型性、平均性は人によつて異なる結果となり易いのである。

そして決定された典型的、平均的な要素の基準に従つて標本構成要素を選定する場合、その規定を完全に満たす要素は容易に見出し得ないために、往々にして観察

し易い要素が選出される結果となり、多くは平均以上（又は以下）の標本が構成されるのが実際である。例えば勤労者の家計調査に於て、被調査者は継続的に克明に日々の家計収支を記帳しなければならぬために、典型的条件よりもこの記帳能力が第一に選択の基準とされ易く、この条件を容易に満たす世帯は消費生活の合理化を心掛けている平均以上の家計であるから、標本は統計集団の平均以上の代表となる傾向が大である。

その上この場合標本より得られた値が、統計集団の悉皆調査の結果とどの程度乖離しているかを測定する客観的な尺度は全く存在しないのである。従つて標本調査の結果の正確性は、標本の構成が典型的な平均的な要素を選んでなされたという確信によるのみであつて、その正確性の程度を客観的に表す事は出来ないものである。

以上有意選択法の正確性を種々考察して来たのであるが、それより明らかた如く有意選択法は、社会的集団現象が異質的な且極めて非類型的な現象より成るた

めに、その典型的、平均的單位の規定並に選択が困難である事にその結果の正確性を欠く最大の理由があるのであつた。従つて社会的集團現象が同質的な且比較的类型的な現象より成る場合には、典型的要素の決定は容易であるため、有意選択法の結果は相当正確であると看做す事が出来るであらう。例えば物価調査に於て調査品目たる商品の、調査日に於けるすべての取引の価格を觀察する事なく、代表的な二三の商店より価格の報告を受けるに止めるのは、価格現象は一物一価の法則より一都市に於ては略々同様の価格が成立する筈であるからである。

### 比例割当法 *quota method*

有意選択により典型的、平均的要素を選定する事の困難を回避するために、典型的單位より成る標本ではなく、統計集團を小規模に再現する如き標本即ち統計集團の完全な縮図であるような標本を構成せんとする方法がある。

### 統計調査法

既に述べたように標本調査法の正確性の基礎は、統計集團の數量的構造と同様の數量的構造を有する標本を構成する事であつた。そこで対照に就て過去の統計調査により統計集團の詳細な構造が判つてゐる時、それによつて統計集團を多くの部分集團に分割し、各部分集團よりその大きさに比例して抽出すべき單位の數を決定し、それを調査員に割当てて各自割当てられた部分 *quota*、即ち割当てられた集團より割当てられた數丈の調査対象を任意に選定、調査せしめる事によつて、対照に就て統計集團の構成割合と一致する如き標本を構成し觀察するのである。此の方法を「比例割当法」といい、輿論調査や価格調査其他の經濟調査によく利用されて來たのである。

例えば一都市の輿論調査の場合、都市人口を対照と考えられる年齢、体性、職業、教育程度等別の人口に細分し、細分された各人口に於て一定の比率で調査すべき單位の數を決定する事によつて、対照に就ての標本及び統計集團の構成比率を同一ならしめるのであ

る。そして各調査員は割当てられた年齢、体性、職業、教育程度の人を適宜探し出して割当数に達する迄調査して来るのであつて、この時調査員は決められた枠の範囲内に於ては可成り大きい選択の自由を持つのである。

しかしながらこの場合割当数の決定に用いる統計が時の遅れによる誤差を有する時は、対照に關して現在の統計集団の正確な縮図たり得ないため、目印に就て統計集団の構造に一致する事は期待し得ず、又調査員に調査対象決定の自由が与えられているため、或る程度調査員自身の便宜によつて選択されるのは不可避であり、調査の困難な対象は敬遠され又調査員の性質によつて対象の質に一定の傾向が現れ易いために、標本の構造は特定の偏りを持つ結果になり勝ちである。そしてこの方法に於ても対照に就て統計集団の構造と同じである時は、目印に就ても又構造は同一であろうという仮定の上に立つているのであるが、その仮定は統計的には必ずしも正しいものではない事は既に考察し

たところである。

#### 標本詳査 *typische Einzelbeobachtung*

標本調査は悉皆調査に比して簡便であり詳細な観察が可能であるにしても、尙統計單位の個別観察、分類、集計の大量観察法の技術的過程を必要とするため、当然に單位の観察の詳細さには限界が存在するのである。従つて社会的集団現象の認識がその要素的現象の性質に就て細部に亘つて注意深い観察を、且長期間継続して行わなければならないような場合には標本調査法によつても把握する事は不可能である。このような場合には社会的集団現象の要素的現象のうちより、典型的、平均的と看做し得るものを極く少数文選択して（尤も社会的集団現象が異質的な要素より成る時は各異質の現象毎に典型的なものを選出する）、それに就て必要な詳細な観察を加えその結果を統計的に整理する事が行われるのであつて、これを「標本詳査」という。

① 従つて此の場合には統計的觀察よりもむしろ個別的觀察

たる性格が濃厚である。

既に述べたように集団構成要素のすべてに對して厳密に典型的、平均的な要素を選択し得るならば、一個の要素の觀察によつても集団全体に就ての性質を知る事が出来るであらう。しかしながら社会的集団現象の場合それは不可能であるため、多数の構成要素を選出して以て典型性を確保しなければならぬのであつた。標本詳査の場合は標本構成要素の数が極めて少いためその觀察結果にはこのような大数法則の作用を予想する事が出来ず、従つて直ちに統計集団（故に社会的集団現象）そのものの性質とは看做し得ないのである。只社会的集団現象が極めて同質的な且比較的類型的な現象より成る場合には、この結果もよく集団全体に妥当し得るであらう。

このように標本詳査は統計集団の悉皆觀察の結果を推定する事は不可能であるが、統計集団の理解に有用な知識を提供するのである。標本詳査は類型的現象の多い自然科学に於て一般に行われる研究方法であるモ

ノグラフィ monographie を社会科学方面に應用したものであつて、ルプレイ Le Play (1803—82) の歐洲労働者の家計調査に始るのである。

## (2) 任意抽出法（抽出調査法）

任意抽出法は統計集団の構成要素のすべてを相互に無關係に（独立に）存在せしめ、且抽出される機会が全く相等的しい状態に置き、それより無作為 *at random* に何等の主観的意図を加えずに構成要素を抽出する事により標本を構成する方法である。

この場合標本構成要素の抽出が無作為であればある程標本より求めた平均値、比例数等は、統計集団の悉皆觀察によつて得られるであらうそれ等の値によく一致し得るのである。蓋し、統計集団構成要素のすべてを相互に独立に且抽出機會の均等な状態に置いて、それより無作為に抽出する時はその抽出行為は全く偶然であるため、或る標識の統計單位が抽出される可能性は、統計集団全体即ち全部集団に於けるその標識の

単位の存在の割合によく比例すると考えられ、従つてかくして抽出せられた統計単位によつて構成される標本は、各標識の統計単位をそれぞれの存在の割合に応じて有する事が期待され、統計集団の縮図たり得る可能性が極めて大であるからである。

任意抽出法によつて構成せられた標本を「任意標本 random sample」といふ、標本を抽出する母体である統計集団を「母集団 population, universe」という。

但しこの場合母集団は正確にいはば諸種の性質を有するものとしての統計集団、即ち種々の標識（性質）を有するものとしての統計単位の集団ではないのであつて、一方向の集団性のみより見た統計集団、故に一の標識のみを有する統計単位の集団であり、従つて統計集団には規定せられた集団性の方向毎に母集団が存在するのである。従つて又一の母集団に対する標本も一標識のみを有するものとしての統計単位の集団である事は明らかであろう。（尚母集団は数学的には統計単位の集団のような実体的な概念ではなく、それ等の統

計単位の有する標識の集合である。従つて又それに対する標本も同じく標識の集合である。）そして母集団は悉皆観察によつて明らかにされ得る構造（分布函数）を有し、その特徴は平均値、比例数等に表され、これ等の値を「母数 parameter」といふ。そして標本より求めた平均値、比例数等は母集団のその推定値であつて、これを「統計量 statistic」といふ。

任意抽出法による時は母集団よりの標本の抽出、構成は全く偶然的な行為となるため、母数とその標本による推定値である統計量との關係が確率論に基いて数理的に規定する事が出来、両者の乖離は偶然的誤差と看做し得るためにその誤差が生ずる可能性はそれによつて数理的に予想し得、従つて標本による推定値の誤差の限界を規定する事が出来るのである。即ち母集団より一定の大きさの標本を可能な限り抽出、構成する時、各標本より求めた統計量の平均が母数に一致する如き統計量（「不偏推定値 unbiased estimate」といふ）の分布状態を表す数式（標本分布函数）を確率

論によつて求める事が出来る。しかれば母集団から任意に抽出せられた一標本より求めた統計量はこの標本分布の一の値をとるのであるから、その出現の可能性

(確率)は標本分布函数によつて表され得るのである。そして統計量の平均が母数に等しいのであるから、母数を中心に正負双方に一定範囲内の値の統計量が出現する確率(信頼度 reliability、その値を信頼係数 confidence coefficient としう)がこれによつて求められ、又反対に統計量の出現の確率が一定の大きさである範囲(信頼区間 confidence interval、これは標本分布函数の 散(標準偏差)で測られる)を知る事が出来る。従つて標本の抽出が全く任意になされる限り抽出された標本より得られた統計量が、標本分布函数より求めた信頼係数が例えば九十五%の信頼区間内にある確率は九十五%であり、故に逆に任意標本より得た統計量にそれを中心に九十五%の信頼区間の誤差を見込む時は、その間に母数を含む確率は九十五%となるのである。かくて一定の信頼度即ち推定値として

の正確性の程度に於て、一定の誤差の範囲を明示して母数を推定する事が出来るのである。

このような標本分布法則は、母集団の構造従つてそれを表す分布函数並に標本構成要素の數及び統計量の如何によつて決るのであるが、それは正規分布、二項分布、Poisson 分布、 $t$  分布、 $\chi^2$  分布、 $F$  分布等として明らかにされているのである。そして相当多くの要素を抽出して標本を構成する時は(即ち大標本の場合)、多くの標本分布法則は正規分布に収斂するのであつて、社会統計調査に於ては抽出數は殆ど大標本の場合であるから、標本分布法則は正規分布をなすものとして誤差を考えればよいのである。

これによつて明らかのように九十五%信頼区間の幅従つて標本分布の分散が小なる程、標本推定値の誤差の幅が小さくなり標本推定は正確となるのである。そして標本分布の分散は標本の大きさが増大するに従つて小さくなる事が数学的に証明されるのであつて、その式より調査目的よりして許容され得る誤差の範囲内の推定値を得るに要する標本の大きさを求める事が出来るのである。

又標本分布の分散は母集団分散の分布（母分散）に正比例するため、母集団が同質的要素より成り各要素の有する値の変動が余り大きくない（故に母分散が小さい）時は、標本分布の分散も小さくなり従つて標本推定の精度は高まるのである。故に母集団が異質的な要素より成りその値の変動が大きい時は、同質的要素の部分集団（層 stratum）に分けて各層内に於ける分散を小さくし、層毎に標本構成要素を抽出する事によつて一定の精度の推定値を得るに必要な標本の大きさ（従つて調査費用）を少くする事が出来、又標本の大きさを一定とすれば誤差の範囲を小さくして推定値の精度を高める事が出来るのである。これを「層化抽出法 stratified sampling」とする。

又母集団が異質的要素より成る時、出来るだけ異質的な要素を集めて部分集団（集落 cluster）を作り、従つて集落間は可能な限り等質的になるようにして母集団を同質的な集落より成る集団に変換し、更に集落の値（即ち集落を構成する統計單位の値の総和）の變

動を小さくするために各集落の大きさを略々同一になるようにし、これより集落を抽出して標本を構成する事によつて同様の効果を挙げる事が出来るであろう。

此の方法を「集落抽出法 cluster sampling」とし、そして集落に母集団構成要素の地域的な集合がとられる場合、これを特に「地域抽出法 area sampling」という。社会的集団現象は普通広大な範囲に亘つて存在しているため、標本構成要素として抽出された要素がばらばらに離れて存在するよりも、出来るだけ地域的に集つている方が個別観察に便利であり且費用も少くて済むであろう。故に地域抽出法が多く用いられるのである。しかしながら實際の調査に於てこのような条件を満す集落を構成する事は困難であり、特に地域抽出の場合には不可能であるため、一般に集落抽出法は精度が落ちるのである。

以上のようにして任意抽出法は標本推定値の誤差を知り得るのみならず、更に一定の精度の推定値を得るに必要な標本の大きさを数理的に規定する事が出来る

のであり、且層化抽出法、集落抽出法による事によつて推定値の精度を高め又は費用を切下げることが出来るために、与えられた費用によつて最大の精度の推定値を得られるように、又要求された精度の推定値を最小の費用で得られるように、標本構成の方法を工夫する事が出来るのである。これを「標本設計 sample design」という。

かくて任意抽出法による時は統計的に正確に統計集団の悉皆観察の結果を推定する事が出来るのであるが、この方法の成否の鍵は標本構成要素の無作為性 randomness がよく確保されているや否やにあるのである。標本構成要素の抽出が無作為になされて始めて、確率数理の基礎の上に展開された母集団分布に対する標本分布法則の適用が可能となるのであり、それによる誤差の算定が有効となるのである。しかしながら統計集団の構成要素のすべてを互に独立に、且等しい抽出機会を持つ状態に置いてそれより無競争に抽出する事は、個々の要素を一連の番号によつて代表せしめ、それよ

り乱数表 table of random sampling の使用等によつて抽出番号を決定して、観察対象を抽出する事によつて容易に実現されるのである。

そして標本構成要素の抽出に際しては抽出対象たり得るすべての要素を完全に且明確に（番号を附し得るように）把握して居らねばならないのであつて、若し抽出対象に脱漏があるならば仮令抽出操作が如何に完全に行われたとしてもその無作為性は損われるのである。故にこの為には統計集団構成要素の悉皆調査が行われねばならないのである。但し実際の調査に於ては他の目的で作成された表を利用し得る場合が多い。（例えば選挙人名簿）又この場合には標本構成要素となるものが抽出操作の結果個々の指定されるのであるため、決定された要素は必ず観察されねばならないのであつて、都合により変更する時はそれだけ抽出の無作為性が害される結果になるのである。

この方法による標本構成は一の偶然的な行為であるため、標本より得られた推定値と母数との差は偶然的

な誤差と看做す事が出来る。そして同じ大いさの標本調査を繰返し行う時、得られる推定値の平均は母数に一致するのであるから任意抽出法は偏りのない推定方法であり、故に任意標本による推定値は偶然的な標本誤差のみを有するのである。（尤も実際の値はこれに大量観察の誤差が附加されるのである。）

先に述べたように任意抽出法に於ては一定の精度の推定値を得るに必要な標本の大いさを求める事が出来るのであるが、それは一母集団即ち統計集団の一方向の集団性に關してのみ妥当する大いさである事に注意しなければならぬ。従つて任意標本調査によつて數個の母数を同時に推定せんとする時は、各母集団毎に要求される精度（目標精度）を実現するに必要な標本構成要素の数を求め、そのうち最大ものを以て標本の大いさを決めねばならないのである。故に任意標本調査は調査項目が多くなると設計が複雑となり技術的困難が増大するため、普通は若干の代表的な項目をとつて設計するか又は調査項目を制限する方法がとられる

のである。悉皆調査の場合はこのような調査項目の制限はなく、幾つでも調査する事が可能であるため、これは任意標本調査法の一の欠点であるといえる。しかしながら統計集団が非常に大きい場合これより多數の要素を抽出して標本を構成する時は、殆どすべての調査項目に就て目標精度を得るに必要な標本の大いさを満す結果になるであらう。従つて目標精度を考慮する事なく、高い抽出率で構成された標本に就て悉皆調査と同様すべての項目を調査する方法が最近行われるようになり、これを *sample-census* という。 *sample-census* は任意標本調査法の調査項目に対する制約を除き、且その技術的簡便性、正確性を利用せんとするものである。我が国に於ては一九五〇年（昭和二十年）世界農業センサスに於て、 $\frac{1}{20}$ の率で抽出された農家に就て詳細な調査が行われ *sample-census* が始めて実施されたのである。

以上任意抽出法による標本調査の特質を種々考察して来たのであるが、それより明らかなく任意標本調

査の実行には、統計集団構成要素のすべてが明確に把握されている事と、標本誤差の推定又は標本の大きさの決定に必要な母集団の分散の値が知られている事が必須条件である。尙更に統計集団の構造に就て詳細な知識があれば、層化法、集落法をとる事によつてより経済的に又はより精度の高い調査が可能となるのであるが、それは任意抽出法によつては心ずしも必要な条件ではないのである。そして統計集団に就ての知識を利用して母集団を層化し集落化するといへ、標本構成要素の抽出は完全な現在の統計集団構成要素の把握に基いてなされるのであるから、仮令その層化、集落化に誤りがあつたにしてもそれは目標精度の達成に影響するのみであつて、標本による推定値を誤らすものではないのである。有意選択法に於ては対照によつて間接的に統計集団の代表たる標本を構成するのであるため、対照と目印との間の仮定の上に、更に対照に就ての統計集団の構造の詳細な知識が決定的な重要性を持つてるのであつて、その正確性は標本調査の結果

果の成否を決するものであつた。しかしながら任意抽出法は直接目印に就て統計集団と等しい値を得られる標本を構成せんとするのであつて、これを可能ならしめるのは統計集団構成要素の無作為抽出であり、それが確保されている限り標本調査結果の正確性は保障されるのである。

#### 抽出集計 sample tabulation

任意抽出法による一部観察による全体の正確な推定は、統計集団の悉皆観察の結果の分類、集計に應用する事が出来る。悉皆調査の場合時の遅れによる統計の誤差を生ぜしめるのは、尨大な統計單位の個別観察の結果の分類、集計の過程である。そこで悉皆観察の結果を母集団としてそれより一部分を無作為抽出して標本を構成し、その分類、集計により全部集計の結果を推定するのである。

この方法を「抽出集計」といい、国勢調査の結果の速報の作成等に利用される。昭和二十五年十月に行わ

れた国勢調査の結果の、十%及び一%の抽出集計による推定が二十六年末に発表された。全部集計の正確な結果は二十八年になつて判る予定である。

× × ×

以上社会的集団現象の統計調査に於てとられる諸種の統計的方法を考察して来たのであるが、それによる社会的集団現象の或る時所に於ける絶対的な大いさと並に構造の正確な認識は、悉皆観察によつて始めて可能となるのである。しかし悉皆調査はそれに要する犠牲が極めて大なる上に、尙且実際上は必ずしも正確な統計が得られないのであつた。そして社会的集団現象の大いさは推算によつて或る程度推定する事が可能であり、又その構造は標本調査によつて推定する事が出来るのである。標本調査は当然誤差を伴うとはいへ悉皆調査に比して著しく費用、労力を軽減し、且統計的観察の誤差を減少せしめるのである。特に任意抽出法による時は標本誤差が推定出来、且その誤差の大い

さを管理する事が出来るために極めて正確な統計的推定が可能であり、場合によつては悉皆調査よりも正確な統計が得られるのである。

かくて従来社会統計調査は一定期間毎の直接的な悉皆調査（即ち Census）を中心に、間接大量観察法による悉皆調査及び有意選択法による標本調査並に推算を併用して行われて来たのであるが、今日では任意抽出法による標本調査法が大いに採入れられ、社会統計調査の体系は悉皆調査と任意標本調査の相互補充による正確な統計を迅速且低廉に作成し得るようによりに再編成されつつあるのである。